EKONOMINIŲ MOKSLŲ BIBLIOTEKA
Nr. 14
Red. *J. Janulaitis*

Albinas Rimka. Statistique. Théorie et méthodologie.

ALBINAS RIMKA

STATISTIKA

TEORIJA IR METODAI

VYTAUTO DIDŽ. UNIVERSITETO TEISIŲ FAKULTETO LEIDINYS
Kaunas A/B „Varpas“ spaustuvė 1939

TURINYS.

Psl.

Turinys V
Lentelių sąrašas X
Autoriaus žodis XII
Literatūra XV

Pirmoji dalis.

A. Statistikos įvadai.

1. Statistikos žodžio kilmė 1
2. Skaičiai, skaičiavimas ir statistika 1
3. Masių vienetai ir statistika 3
4. Statistikos mokslas ir gyvenimas 4
5. Statistikos dėsningumai 7

B. Statistikos surašinėjimų istorinė apžvalga.

6. Surašinėjimai ir statistika senovės gadynėj 9
7. Viduramžių surašinėjimai 12
8. Lietuvos surašinėjimai ir 1528 m. „popis zemsky“ . . . 14
9. Naujųjų laikų statistika 15

C. Statistikos mokslo istorinė apžvalga.

10. Statistika senovėje ir viduramžiais 17
11. Renessansas ir vad. universitetinės statistikos kilimas . . . 17
12. Achenwall'io „statistika“ 19
13. Universitetinės statistikos pasekėjai ir jų ginčai 21
14. Graunt'o „Natūra“ ir Suessmilch'o „Dieviškoji santvarka“ . . 22
15. W. Petty politinė aritmetika 24
16. Quetelet ir jo „Socialinė fizika“ 25
17. Pagrindinės Quetelet mokslo mintys 29
18. Moderniosios statistikos krypsniai 30

D. Statistikos darbo organizacija.

19. Statistikos darbo organizacijos principai 32
20. Statistikos darbo organizacija Lietuvoje: 34
    a) Statistikos darbo organizacija prieš D. karą . . . 34
    b) St. darbas po 1918 m. 35
    c) St. darbo organizacija Klaipėdos krašte . . . 39
    d) St. darbo reorganizacija 1930 m. 43
21. Tarptautinės statistikos organizacija 46
22. Statistikos šaltiniai 48
    a) Statistikos šaltinių charakteristika 48
    b) Lietuvos statistikos šaltiniai 49
    c) Bendrieji tarptautinės statistikos šaltiniai 50
23. Statistikos vadovėliai ir periodinė literatūra: 51
    a) Populiaresnieji st. teorijos vadovėliai 51
    b) Matematinės statistikos vadovėliai ir vadovėliai su matematinio analizio metodų nuodugnesniu aiškinimu . . 52
    c) Periodinė st. literatūra 52

Antroji dalis.

A. Bendrieji statistikos teorijos ir metodologijos klausimai.

Psl.

1. Statistikos metodų pagrindinės ypatybės 53
2. Kai kurios statistinės sąvokos ir terminai 56
3. Statistikos darbo etapai ir jų uždaviniai 58

B. Statistinio stebėjimo pagrindinės sąlygos.

4. Masių stovio arba padėties statistinis stebėjimas. 59
   a) Skaitomoji diena stovio statistiniam stebėjimui . . 59
   b) Stebėjimo kritiškasis momentas 60
   c) Surašomoji diena arba surašomasis laikas 61
5. Judėjimo arba reiškinių bei atsitikimų statistinis stebėjimas . 62
   a) Skaitomasis periodas 62
   b) Kritiškasis momentas 64
   c) Stebimasis arba surašomasis laikas 65

C. Statistinio stebėjimo metodai ir sistemos.

6. Bendrosios pastabos 66
7. Tiesioginis ir netiesioginis st. stebėjimas: 67
   a) Tiesioginė statistika 68
   b) Netiesioginė statistika 68
   c) Pranašumo bei pirmenybės klausimas 69
   d) Tiesioginės ir netiesioginės st. metodų kombinacija arba junginys 72
8. Individualinis ir kolektyvinis st. stebėjimas: 73
   a) Individualinis st. stebėjimas 73
   b) Kolektyvinis st. stebėjimas 74
   c) Pranašumo arba pirmenybės klausimas 75
9. Ištisinis ir dalinis st. stebėjimas bei reprezentacinis metodas: . 77
   a) Ištisinis st. stebėjimas 77
   b) Dalinis st. stebėjimas 77
   c) Reprezentacinis metodas statistikoj 78
   d) Atsitiktinių reprezentantų st. stebėjimas 79
   e) Tipiškųjų reprezentantų parinkimas 80
   f) Pranašumo bei pirmenybės klausimas 81
   g) Elementarinis reprezentacinio metodo pagrindimas . . 83
10. Statistinio stebėjimo organizavimas: 85
   a) Bendrosios žinios 85
   b) Surašomojo lakšto paruošimas 86
   c) Statistinės klasifikacijos planas 86
   d) Statistiškai prasmingo surašomojo lakšto ypatybės . . 87
   e) Statistikos stebėjimo kontrolė 88
11. Priedėliai: 90
   a) Surašomasis lakštas individualiniam stebėjimui kortelių sistema 90
   b) Surašomasis lakštas individualiniam st. stebėjimui sąrašų sistema 92
   c) Surašomojo lakšto pavyzdys kolektyviniam st. stebėjimui . 93
   d) Pavyzdys skaitomosios kortelės, naudojamos mechaniškam skaitomųjų žymių ir pažymių apdirbimui (klasifikacijai) Hollerith mašinų sistemoj 94

Trečioji dalis.

A. Statistikos vaizdavimas bei klasifikacija.

Psl.

1. Statistinio vaizdavimo esmė ir uždaviniai 95
2. Statistinės klasifikacijos rūšys ir jų metodologinis pagrindimas . 96
3. Kokybinė klasifikacija ir jos uždaviniai: 98
   a) Statistinis tikslingumas bei prasmingumas kokybinėj klasifikacijoj 98
   b) Statistiškai prasmingos klasifikacijos sąlygos . . . 100
4. Kiekybinė klasifikacija (dydžio grupių bei klasių sudarymas): . 101
   a) Kiekybinės klasifikacijos tikslas ir bendrieji dėsniai . . 101
   b) Statistinės klasifikacijos pagrindai 103
   c) Aukštesniojo laipsnio kiekybinės klasės 106
   d) Kiekybinė klasifikacija Galton'o metodu 107
   e) Kumuliatyvinis metodas kiekybinėj klasifikacijoj . . . 108

B. Statistinės klasifikacijos sistemos.

5. Statistikos lentelės ir terminai: 109
   a) Statistikos lentelių sudarymas ir kai kurie st. technikos terminai 109
   b) Intervalo apibrėžimas kiekybinėj klasifikacijoj . . . 113
6. Statistikos lentelių sudarymo pavyzdžiai: 115
   a) Paprastos esminėmis žymėmis sudarytos lentelės . . 115
   b) Kiekybiniais pažymiais sudarytos st. lentelės . . . 116
   c) Klaidingos st. klasifikacijos st. lentelių pavyzdžiai . . 119
7. Statistikos duomenų vaizdavimas grafikos priemonėmis . . 123

Ketvirtoji dalis.

A. St. eilučių santykis ir statistikos analizio metodai.

1. Statistinio analizio formos ir priemonės 133
2. Statistikos eilučių santykį matuojamieji skaičiai 134
3. Statistiniai santykio daliniai 135
4. Statistinio santykio lyginiai: 137
   a) Bendrieji (brutto) statistinio santykio lyginiai . . . 143
   b) Grynieji (netto) statistinio santykio lyginiai . . . 145
   c) Specifiniai statistinio santykio lyginiai 154

B. Vidurkiai ir st. eilučių dispersija.

5. Statistikos vidurkiai: 157
   a) Aritmetinių vidurkių apibendrinimai bei formulos . . 159
   b) Vidurkiai ir jų suradimą palengvinantieji sk. metodai . . 160
   c) Aritmetinio vidurkio skaičiavimo pavyzdžiai . . . 163
   d) Vidurkiai iš dydžio grupių bei klasių 169
   e) Medianas 172
   f) Dominantas 176
   g) Matematiniai vidurkiai 182
   h) Vidurkių statistiniai uždaviniai ir jų tikslingas parinkimas . 185
6. Statistikos eilučių dydžių pasiskaidymas bei dispersija: . . 192
   a) St. dispersijos sąvoka ir rūšys (tipai) 192

Psl.

b) Kvartiliai ir kvartiliniai koeficientai 197
c) Paprastieji arba aritmetiniai nukrypimai bei asimetrijos ir dispersijos koeficientai 204
d) Normalioji dispersija ir vidurinis kvadratingasis nukrypimas 208
e) Apytikriai skaičiuojamas metodas kvadratingiems nukrypimams surasti 225

C. Statistikos dėsningumai ir tikimybės.

7. Statistinės ir matematinės tikimybės: 234
a) Vad. statistikos dėsningumai ir kai kurie jų pasireiškimai 234
b) Atsitiktinių įvykių eigos dėsningumai 240
c) Atsitiktinių įvykių matematiškoji tikimybė 244
8. Patikimoji, vidutinė (normali) ir galimoji paklaida statistikos skaičių eilutėse 258
9. Tikimybės ir atsitiktinės paklaidos reprezentacinėj statistikoj: 269
a) Bendros pastabos 269
b) Reprezentacinės statistikos dalinių tikimybė bei paklaidos 273
c) Reprezentacinių st. vidurkių tikimybė bei paklaidos . . 279
d) Orientacinių bei „geriausiųjų" statistinės masės reprezentantų parinkimas 281
10. Interpoliacija 290
a) Interpoliacijos esmė 290
b) Interpoliacijos metodai 292
1) Grafiškoji bei aritmetinė interpoliacija tiesės lygtimi 292
2) Geometrinė interpoliacija pagal sudėtinių procentų formulą 294
3) Interpoliacija vidurkiais 298
4) Kiti interpoliacijos metodai 299
5) Linkmės dalinių interpoliacija 300
6) Linkmės dalinių ekstrapoliacija 303
11. Koreliacija ir statistinės pareinamybės: 304
a) Koreliacijos sąvoka ir rūšys 304
b) Koreliacijos koeficiento pakaitalai 305
c) Pearson'o koreliacijos koeficientai 308
d) Kolektyvinis koreliacijos koeficientas 315
12. Linkmės kitimų tikimybė ir hipotetiniai dydžiai 319
a) Linkmės kitimų statistinė prigimtis 319
b) Linkmės kitimų hipotetiniai dydžiai 321

D. Statistikos eilučių išlyginimai ir pagrindinės linkmės apibrėžimai.

13. St. eilučių išlyginimų esmė ir statistinė prasmė 326
14. Statistikos eilučių pagrindinės linkmės apibrėžimas . . . 328
a) Mažiausiųjų kvadratų metodas pagrindinei linkmei pabrėžti 329
b) St. eilučių pirmaeilė linkmė arba linkmės kitimų tiesė . 331
c) Statistikos eilučių antraeilė, trečiaeilė ir t. t. linkmė . . 337
d) St. eilučių išlyginimas bei pagrindinės linkmės apibrėžimas kintamaisiais vidurkiais 338
15. Kintamųjų parametrų metodas st. eilutėms išlyginti . . . 341
16. Sumavimų metodas statistikos eilutėms išlyginti bei kitimų linkmei aprėžti 347
17. Statistinių eilučių linkmės bei išlygintųjų dydžių patikimumo patikrinimas 358

E. Statistikos indeksai.

Psl.

18. Statistikos indeksų esmė ir rūšys: 361
19. Indeksai su pastovia baze: 362
    a) Indeksai su baze iš aritmetinio vidurkio 363
    b) Indeksai su baze iš mediano 364
    c) Indeksai su baze iš dominanto 364
    d) Indeksai su baze iš „norminių" dydžių 365
    e) Indeksai su baze iš kraštutinių dydžių 366
    f) Indeksai su baze iš specialiai parinktų dydžių . . . 366
20. Indeksai su nepastovia (keičiama) baze 367
    a) Kintamieji indeksai 368
    b) Jungtiniai (kumuliatyviniai) indeksai 369
    c) Periodiniai bei sezoniniai indeksai 369
        1) Paprastųjų periodų indeksai 370
        2) Aprėžtųjų periodų indeksai 371
        3) Normalūs periodiniai bei sezoniniai indeksai . . 373
    d) Periodinių vidurkių metodas normaliai sezoniniams indeksams skaičiuoti 374
    e) Periodiškai kolektyvinių indeksų metodas 376
    f) Persons'o metodas 379
21. Kolektyviniai indeksai 384
    a) Kolektyvinių indeksų esmė ir jų sudarymo sąlygos . . 384
    b) Kolektyvinių indeksų sudarymo metodai 385
        a) Paprastieji aritmetiniai indeksai 385
        b) Svertiniai indeksai 386
    c) Svorio koeficientų indeksams parinkimo sąlygos . . 386
    d) Individualinių indeksų svėrimo sistemos 389
    e) Kolektyvinių indeksų svorio koeficientų apibrėžimas . 392
    f) Matematiniai indeksai 398
    g) Irving Fischer'io „ideališkoji" indeksavimo formula . 401
    h) Skirtingais metodais gautųjų kolektyvinių indeksų palyginimas 403

Priedėliai 405

Vardų rodyklė 413

Žodžių ir dalykų rodyklė 414

Sommaire 419

Lentelių sąrašas.

Eil. Nr.		Psl.
1.	Natūrinis gyventojų judėjimas 1930—1936 m.	115
1a.	Gimusieji 1936 m.	115
2.	Gimusieji 1934—1936 m.	116
3.	Mažeikių, Šakių ir Ukmergės žemės ūkiai, sudoroti į dydžio grupes Galton'o metodu	118
4.	Kumuliatyvinių dydžių grupių sudarymo schema	119
5.	Klaidinanti lentelė apie Lietuvos gyventojus tautybėmis	120
6.	Susisiekimo priemonės Kaune 1919—1934 m.	121
7.	Žaliakalnio keltuvo keleivių skaičius 1934 m.	121
8.	Miesto butų kolonijų gyventojai 1934 metais	122
9.	Lietuvos gyventojai 1924 m. pradžioje tikybomis	135
10.	Gimusiųjų santykio skaičiai	138
11.	Lietuvos gyventojai 1924 m. ir mirusieji amžiaus grupėmis	156
12.	Lietuvos gyventojai 1924 m. ir mirusieji amžiaus grupėmis	156
13.	Mirimai Lietuvoje 1926—1935 m.	163
14.	Lietuvos lauko darbininkų dienos uždarbis 1924 m. pavasarį	166
15.	Vidutinio uždarbio (svertinio vidurkio) skaičiavimo schema sumavimų metodu	167
16.	Lietuvos žemės ūkiai dydžio klasėmis	170
17a.	Žieminių kviečių derliai 1932—34 m.	180
17.	Kviečių derliai iš 1 ha 1932—34 m.	207
18.	Galimų a ir b įvykių skaičius lygiose sąlygose	214
19.	Binomo bei jo koeficientų lentelė $(a+b)^n=2^{10}$	215
20.	Binomo bei jo koeficientų lentelė, kur $(a+b)^n=2^{14}$	216
21.	Kviečių derliai 1932—34 m.	219
22.	Lietuvos žemės ūkių kvadratingajam nukrypimui skaičiuoti schema	221
23.	XXVII formulos a ir b varijantų pritaikomojo skaičiavimo schema	225
24.	1932—24 m. kviečių derliai Lietuvoje ir jų kvadratingieji nukrypimai	227
25.	1932—34 m. kviečių derliai Lietuvoje ir jų kvadratingieji nukrypimai	228
26.	Kvadratingieji nukrypimai iš nuotolinių eilutės Nr.	229
27.	Schematinė lentelė viduriniam kvadratingajam nukrypimui sumavimų metodu surasti	231
28.	Lietuvoje 1000-čiui gyventojų tenka	235
29a.	Mėnesiniai jungtuvių skaičiai	236
29b.	Mėnesiniai mirimų skaičiai	237
30.	Gimimų skaičiai Lietuvoje	238
30a.	Westergaard'o tyrinėtųjų dviejų vienodai galimų įvykių atsitiktiniai nuo „normos“ nukrypimai, kai „normos“ b ir r santykiauja kaip 50:50	242
31.	Žieminių kviečių derliai 1932—34 m.	256
32.	Lietuvos gyventojų skaičiaus tiesinė (aritmetinė) interpoliacija 1924—1927 metams (1000 gyv.)	293
33.	Aritmetinės ir geometrinės interpoliacijos pavyzdys	296
34.	Geometrinės interpoliacijos pavyzdys logaritmų pagalba	297
35.	Linkmės dalinių interpoliacija pasikliaujant koeficientų dispersijos norma	302

Eil. Nr. Psl.
36. Gimimų ir natūrinio prieauglio „pareinamybės indeksai“ pagal Julin'ą 306
37. Gimimų ir natūrinio prieauglio koreliacijos koeficientas sulig skirtumais nuo vidurkio 310
38. Gimimų ir natūrinio prieauglio koreliacijos koeficientas sulig pirmaisiais skirtumais 314
39. Schematiniai duomenys gimimų, mirimų ir natūrinio prieauglio koreliacijos koeficientams 317
40. Pavyzdinio skaičiavimo schema natūrinio prieauglio hipotetiniams dydžiams 324
41. Natūrinio prieauglio linkmės kitimų tiesės skaičiavimo schema normalių regresijos lygčių pagalba 332
42. Linkmės kitimų tiesės skaičiavimo schema mažiausiųjų kvadratų metodu, kai n yra nelyginis skaičius 335
43. Linkmės kitimų tiesės skaičiavimo schema mažiausiųjų kvadratų metodu, kai n yra lyginis skaičius 336
44. St. eilučių išlyginimo schema kintamaisiais vidurkiais . . 342
45. Kreivės ordinatų dispersijos apibrėžimas trejetuko kintamaisiais intervalais 342
46. Sumavimų metodo schema natūrinio gyventojų prieauglio duomenimis 1922—37 m. 353
47. Sumavimų metodų gautieji tiesinės (y_1) ir antraeilės (y_2), trečiaeilės (y_3) bei ketvirtaeilės (y_4) linkmės duomenys 356
48. Skaičiavimo schematinis pavyzdys išlygintųjų st. eilučių skirtuminei dispersijai apibrėžti. 359
49. Periodinių vidurkių metodo schema 375
50. Normaliai sezoninis indeksas periodinių vidurkių metodu (aritmetinis) 377
51. Normaliai sezoninis indeksas periodiškai kolektyvinių indeksų (periodografiniu) metodu 378
52. Normaliai sezoninis indeksas Persons'o metodu 381
53. Normaliai sezoninių indeksų palyginimas ir periodinių vidurkių indekso sezoninių svyravimų eliminavimas 382
54a. Kainos ir jų svorio koeficientai 392
54b. Svertinių indeksų skaičiavimo schema betarpiškai iš kainų kitimo pagal lentelės 54a duomenis 393
54c. Aritmetiniai ir svertiniai kolektyviniai indeksai iš individualinių indeksų 394
55. Prekių grupių kainų indeksai (1926—29 m.=100) 397

Autoriaus žodis.

Nuo 1924 mt., kai parašiau „Statistikos pagrindus", statistikos teorija ir praktika pažengė toli į priekį. Šių dienų socialekonomika (pav., ūkio konjunktūros tyrimai) ir demografija (ypač natūrinis gyventojų judėjimas) statistikai stato visą eilę problemų, kurių sekmingas išnagrinėjimas tegalimas pasikliaujant griežtesniais matematinio analizio metodais. Dėlto jau 1931 mt. išleistoji „Socialekonominė statistika" kai kuriuos matematinio analizio metodus trumpai apibrėžė, o paskaitose, pratybose bei statistikos seminaro darbuose tuos metodus tekdavo vis daugiau ir daugiau studijų objektu daryti. Tačiau atitinkamo vadovėlio stoka statistikos studijų eigą žymiai trukdydavo, nes 1924 mt. išleistuose „Statistikos pagrinduose" daugelis klausimų nepaliesta, o kiti nepakankamai išnagrinėti.

Dabar leidžiamoji „Statistika" yra keliolikos metų su studentais bendradarbiavimo vaisius. Mūsų Teisių Fakulteto studentų auditorija statistikos studijų atžvilgiu yra gana marga. Pirmiausia, studentai teisininkai į statistiką žiūri kaip į antraeilį bendrojo išsimokslinimo dalyką. Jie reikalingi bendrųjų žinių iš demografinės, vad. intelektualinės bei kriminalinės statistikos, taip pat privalo pažinti tų sričių statistikos šaltinius ir sugebėti kritiškai įvertinti statistikos duomenimis pagrindžiamą argumentaciją. Bet kritiškas statistikos duomenų ir statistinės argumentacijos vertinimas nėra galimas be statistikos teorijos bei metodų pažinimo. O modernioji statistikos teorija bei metodologija negali išvengti matematikos tiekiamųjų analizio priemonių bei... tikimybių, kurių, deja, vis dar bereikalingai bijo ne vieni studentai. Dėl šios priežasties ekonomistai dar prieš statistikos studijas specialiai supažindinami su binomu ir kitais vidurinių mokyklų matematikos programose esamais, bet pirmojo kurso studentų galvose retai tesutinkamais, dalykais. Be to, šių dienų ekonomistui statistika yra vienas iš pačių pagrindinių mokslo dalykų, nes daugelis valstybės bei savivaldybinių

įstaigų, taip pat stambesnės verslo įmonės, yra reikalingos žmonių, kurie galėtų savarankiškai statistikos darbus dirbti, o reikale ir juos organizuoti.

Tuo būdu statistikos kursą tenka dėstyti žmonėms su nevienodu pasiruošimu ir su skirtingais reikalavimais. Ši aplinkybė yra paveikusi „Statistikos" turinį ir net patį planą. Statistikoj bendro išmokslinimo žinių ieškantiems teko platokai atpasakoti (iš „Statistikos pagrindų") statistikos istoriją, patiekti visą eilę statistikos duomenų, lentelių, net įvairių formaliai juridinio pobūdžio nuostatų ir kt. Iš kitos pusės, „Statistikoj" teko aptarti net tokius klausimus, kurie teorijos paskaitų bendrame kurse nenagrinėjami ar tik labai trumpai paliečiami, nes jie sudaro studijų objektą tik statistikos seminaro dalyviams. Taigi „Statistika" apima tiek dalykus, su statistikos teorija bei metodologija mažai tesusietus, tiek dalykus, prašokančius Teisių Fakulteto studentams bendrai privalomo kurso programą. Tačiau šitą „Statistikos" „apsunkinimą" padariau sąmoningai šiais sumetimais:

1. teoretinės paskaitos studijuojantiems statistiką tėra tik įvedamieji paaiškinimai pratybų darbams;
2. ne visi studentai išgali statistikos paskaitas reguliariai lankyti;
3. su statistikos teorijos ir metodologijos klausimais šiandien tenka susidurti ir tiems, kurie universiteto lankyti negali.

Taigi „Statistiką" rašydamas turėjau galvoje ne tik akademinio jaunimo studijas, bet ir savimoką. Ši aplinkybė iš dalies nulėmė ir „Statistikos" planą bei atskirų dalių proporciją. Pav., statistikos analizio metodams tik dėl to tiek daug vietos pašvęsta, kad turėta prieš akis ir tie „Statistikos" skaitytojai, kurie nei paskaitų, nei pratybų, nei statistikos seminaro lankyti negali.

Kiek šis darbas pavyko — patsai nedrįstu spręsti. Galiu tik pasakyti, kad pati „Statistika" jau „galutinai" buvo parašyta prieš penketą metų, paskui tris kartus iš naujo viskas pertvarkyta ir tik šiais mokslo metais jau nekeista. Visi šie keitimai buvo daromi pasirėmus paskaitų, pratybų ir statistikos seminaro darbų patyrimais. Ypač su statistikos pratybų

ir seminaro dalyviais bendradarbiavimas bei įvairūs statistikos teorijos bei metodologijos klausimų aptarimai daug prisidėjo prie „Statistikos“ turinio, plano ir net kai kurių terminų apibrėžimo. Šia proga ir noriu visiems stitistikos pratybų ir seminaro dalyviams savo padėkos žodžius pareikšti. Ypatinga padėka priklauso Dr. D. Budriui ir Dr. J. Bučui, kurie jau eilė metų studentų statistikos pratybose vadovaudami savo patyrimu paremtų pastabų suteikia, ir p. p. J. Žukauskui, K. Palčiauskui, J. Meškui ir kt., kurie statistikos seminare keletą metų kaip savanoriai priv. asistentai aktingai bendradarbiaudami daug prisidėjo ir prie „Statistikoj“ nagrinėjamų klausimų išryškinimo.

Visai atskirai noriu išreikšti savo nuoširdžią padėką prof. Viktorui Biržiškai, kuris nepagailėjo laiko matematinio analizio skyrius peržiūrėti ir kurio kritikos pastabos man ir visam darbui buvo itin naudingos. Taip pat esu dėkingas statistikos seminaro dalyviui, diplomuotam matematikui ir ekonomikos studentui A. Jurskiui už vertingas pastabas, kurias padarė IV dalies (stat. matematinio analizio) korektūras skaitydamas. Be to, Dr. Dz. Budrys ištisai visos knygos korektūras skaitė ir vardų bei žodžių (terminų) rodyklę sudarė, o p. Rudaitytė-Šiurkienė daug darbo įdėjo perrašinėdama, skaičiuodama bei skaičiavimo rezultatus tikrindama, taip pat skaitydama korektūras ir kt. Visų čia minėtų asmenų darbas ar žodis daug pagelbėjo ir palengvino dirbti, ir bent šioj vietoj lai bus leista pareikšti gilios padėkos žodį.

1939 m. gegužės 3 d. *A. Rimka.*

LITERATŪRA.

Aftalion, A.: Cours de Statistique. 3-e Édition. Paris.

Ancyferov, A. N.: Kurs statistiki. Praga, 1922.

Anderson, O.: Einführung in die Mathematische Statistik, Wien, 1935.

Bertillon, J.: Cours Elementaire de Statistique Administrative. Paris.

Bojarsky, Brand ir kt.: Statistika. Maskva, 1936.

Bowley, A. L.: Elements of Statistics. London, 1920.

Bowley, A. L.: The Nature and Purpose of the Measurment of Social Phenomena. London, 1923.

Connor, L. R.: Statistics in Theory and Practice. London, 1932.

Conrad-Hesse: Allgemeine Statistik. Jena, 1923.

Czuber-Burkhardt: Die statistischen Forschungsmethoden. Wien, 1938.

Donner, Otto: Statistik, Hamburg. 1938.

Ellis, Henry: A General Introduction to the Domesday Book. London, 1933.

Fisher, Irving: The Making of Index Numbers. Boston and New York, 1923.

Fisher, A.: The Methematical Theory of Probabilities. New York, 1936.

Flaskämper, P.: Theorie der Indexzahlen. Berlin, 1928.

Florence, Sargant P.: The Statistical Method in Economics and Political Science. London, 1929.

Harper, F. H.: Elements of Practical Statistics. New York, 1930.

Hull, C. H.: The Economic Writings of Sir William Petty. Cambridge. 1899.

Jerneman, Tor: On the Substitution of Constans in Sampling Formulae with special Reference to the Theorie of the „too small values". — Nordic Statistical Journal, Vol. 3, Stockholm, 1931.

Julin, A.: Principes de Statistique Theoretique et Appliquée. I. Statistique Theoretique. Paris-Bruxelles.

Yule, G. U.: An Introduction to the Theory of Statistics. London, 1922.

Kaufman, A.: Teorija i Metody Statistiki.

King, W. I.: Elements of Statistical Method. 1923.

Liubavskij: Oblastnoje delenije i miestn. uprav. Litovsko-Ruskavo Gasudarstva. Maskva. 1892.

Lorenz, P.: Der Trend. Berlin, 1931. (Vierteljahrshefte zur Konjunkturforschung, Sonderheft 21).

Mayer, G. von. Statistik und Gesellschaftslehre. I. Theoretische Statistik. Tübingen, 1914.

March, L.: Les Principes de la Methode Statistique. Paris, 1930.
Moncetz, A.: Initiation aux Methodes de la Statistique. Paris, 1935.
Müller, J.: Theorie und Technik der Statistik. Jena, 1927.
Ping-Hua Lee, Mabel: The Economic History of China. New York, 1931.
Quetelet, L. A. J.: Sur l'home et le développement de ses facultés. Pirmasis leidimas išėjo Paryžiuje 1835.
Rietz-Bauer: Handbuch der mathematischen Statistik. Leipzig und Berlin, 1930.
Rimka, A.: Socialekonominė Statistika. Kaunas, 1933.
Rimka, A.: Statistikos Pagrindai. Kaunas, 1925.
Rimka, A.: Lietuvos visuomenės ūkio bruožai. 1924.
Romanovsky: Elementarnyj Kurs Matematičeskoj Statistiki.
Schott, S.: Statistik. Berlin, 1920.
Schwartz: Philosophie der Statistik. — Allgemeines Statistisches Archiv, 21. Bd., II. Heft, 1931.
Suessmilch, J. P.: Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts aus der Geburt, dem Tode und der Fortpflanzung. Berlin, 1741.
Tarulis, A.: Didž. Lietuvos Kunigaikštijos karo prievolių (1528 m.) sąrašų patikimumas statistikai. — Ekonomika 2, 1936.
Vinci, F.: Manuale di Statistica. Bologna, 1936.
Wagemann, E.: Narrenspiegel der Statistik. Hamburg, 1935.
Wagenführ, R.: Statistik leicht gemacht. Hamburg, 1934.
Westergaard und Nybolle: Grungzüge der Theorie der Statistik. Jena, 1928.
Willford, J. King: The Elements of Statistical Method. London.
Winkler, W.: Grundriss der Statistik. Berlin, 1931.
Winkler, W.: Die statistischen Verhältniszahlen. Leipzig und Wien, 1923.
Winkler, W.: Statistik. Leipzig, 1933.
Winkler, W.: Das Problem der Willensfreiheit in der Statistik. — Revue de l'Institut International de Statistique, Livre 2, 1937.
Winkler, W.: Einige alte und neue Masse des natürlichen Bevölkerungswachstums. — Revue de l'Institut International de Statistique, Livre 1, 1938.
Žižek, F.: Grundriss der Statistik. München und Leipzig, 1921.
Žižek, Fr. : Die statistischen Mittelwerte. Leipzig, 1908.
Žižek, F.: Nichtkorrekte statistische Verfahren. — Allg. Statistisches Archiv, 21. Bd., I. Heft, 1931.

P. S. Be to, žiūr. literatūros ir šaltinių sąrašą „Socialekonominėj Statistikoj" (p. p. V—VIII).

PIRMOJI DALIS.

A. Statistikos įvadai.

1. Statistikos žodžio kilmė.

Žodis „statistika“ šiandien girdimas kiekviename žingsnyje. Štai, pav., rado skaitytojas laikraštyje porą skaičių apie gimimus ar užsienių prekybą — jau taria, kad tai „statistika“. Suskaitė įstaigos viršininkas tarnautojus arba mokytojas savo mokinius — ir vėl „statistika“. Pamatė studentas knygą su daugeliu skaičių arba st. lentelių — taip pat „statistika“. Žodžiu, kur tik pasisuksi — vis rasi „statistiką“. Šiandien kalbama ne tik apie gyventojų statistiką, bet ir apie tautų, tikybų, namų, gyvulių, žemės ir kt. daiktų statistiką, taip pat apie mirtingumo, prasikalstamumo, konjunktūros ir kt. statistiką. O neseni dar laikai, kai ir žodžio—„statistika“—niekas nepažino, arba tuo žodžiu visai ką kita vadino. Dar XVIII šimtm. ginčytasi, ar žodis „statistika“ yra kilęs iš s t a t o — valstybė, ar iš s t a t u s — stovis, padėtis. Šis ginčas taip ir pasibaigė statistikos žodžio kilmės neišsprendęs. Bet tai ir nesvarbu. Svarbiau yra šio žodžio šiandieninę prasmę gerai suprasti.

2. Skaičiai, skaičiavimas ir statistika.

Visų pirma statistika siejama su skaičiais ir skaičiavimu. Bet ne kiekvienas skaičius yra statistikos skaičius ir ne kiekvienas skaičiavimas yra statistika. Jei autobuso konduktorius suskaito keleivius arba piemenukas — savo ganomas avis, tai dar nėra statistika. Nebus statistika ir bet kurie skaičiai, jei jie reikalingos formos neturi. Pav., 2,5 mil. gyventojų arba 300.000 ūkių yra tos pačios rūšies informaciniai skaičiai, kaip ir 100 knygų, 10 ligonių, 5 gaisrai ir t. t. Bet tie patys skaičiai gali būti ir statistikos skaičiais.

Skaičių, skaičiavimo ir statistikos savitarpio ryšį ir santykius vaizdingai paaiškina S c h o t t'as dviem pavyzdžiais[1]. Schottas vienoj pusėj rodo mergaitę mezgėją, kuri iš nuobodumo skaičiuoja mezgamąsias akis, o antrojoj — Don Juano tarną Leporello, kuris, paėmęs savo pono apleistųjų mergaičių sąrašą, nustebusiai donnai Elvyrai išdėsto: Italijoj buvę jų 640, Vokietijoj — 231, Prancūzijoj — 100, Turkijoj — 91, o gerbiamosios donnos Elvyros tėvynėj Ispanijoj — 1003. Mezgėja mergaitė, anot Schotto, nors milijonus mezgamųjų akių suskaitytų, nieko bendra su statistika neturi; tuo tarpu Leporello skaičiai jau bus statistika ir jis pats statistikos darbą atliko. Skirtumas gi tarp mezgėjos ir Don Juano tarno tik tas, kad anoji vien skaičiavo ir skaičiavo, o Leporello, vesdamas savo pono mylėtų mergelių sąrašą, suskirstė jas atsižvelgdamas tam tikrų ypatybių (šiuo atveju — pilietybės pažymio).

Taigi s t a t i s t i k a nėra paprastas suskaitymas nei skaičiavimas, o s i s t e m i n g a s i r p r a s m i n g a s ž y m ė j i m a s s k a i č i a i s a t s i ž v e l g i a n t t a m t i k r ų ž y m i m o j o d a i k t o a r b a d a l y k o y p a t y b i ų. Vadinasi, patys vieni Lietuvos gyvenamųjų vietovių arba kurios nors vietovės gyventojų sąrašai, arba kad ir sistematinis bibliotekos knygų katalogas, daromas žiūrint turinio, autorių arba išleidimo metų, nebus statistika. Kurio nors sodžiaus, valsčiaus arba miesto gyventojų sąrašas tik tuomet galės būti vadinamas statistika, kai surašytieji gyventojai bus suskirstyti ne atsitiktinai, bet tam tikrais ir aiškiai aprėžtais pažymiais, k. a.: amžius, tikyba, tautybė, lytis, verslas, turto ar pajamų dydis ir t. t. Taip pat gyvenamųjų vietovių statistika bus tik tada, kai kalbamos vietovės bus suskirstytos pagal geografinę padėtį, viešosios teisės santykius ar kt. Tas pats ir su knygų katalogu: reikia tik bibliotekos knygų katalogą paėmus kiekvieno, pagal turinį ar kt. sudaryto, skyriaus knygų skaičius surašyti eilutės sekoj, ir jau turėsime knygyno knygų statistiką skyriais; o tų pačių knygų skaičiai, suskirstyti autoriais, leidėjais arba išleidimo metais, duos kalbamų knygų leidėjų, autorių arba išleidimo metų statistiką. Ir visuotinis visos valstybės gyventojų surašymas savaime duoda tik medžiagą statistikai. To

[1] *Schott,* Statistik, psl. 6/7.

surašymo gyventojų statistiką turėsime tiktai tada, kai surašymo duomenys bus suskirstyti tam tikromis aiškiai apibrėžtomis ypatybėmis (pav., lyties, tautybės, tikybos, verslo, šeimos, turto sudėties ar kt. pažymiais).

3. Masių vienetai ir statistika.

Bet ne tik forma ar tam tikra žymėjimo sistema nulemia skaičių „statistiškumą“. Mat, statistikai turi reikšmės ne vienetai ir ne atskirų vienetų individualinės, nuo kitų visai skirtingos, ypatybės, bet ištisos masės, arba kurie nors masiniai reiškiniai. Statistikai reikšmės turi ir atskirų vienetų ypatybės, — bet tik tokios, kurios sutinkamos ir kituose, nors ir ne visuose, masės vienetuose. Jei, pavyzdžiui, paimti kurios nors valstybės ar krašto gyventojų surašymo duomenis, tai ten ministeri, prezidentą ir karalių atstovauja lygiai toks pat skaitmeninis vienetas, kaip ir elgetą ar sodžiaus piemenuką. Tas pats ir mirimų, gimimų ir kt. statistikoj. Net suskirsčius gyventojus verslo ar turto pažymiais, niekuomet nežiūrima atskirų asmenų verslo ar turto, o žiūrima tik ištisų grupių ar sluoksnių, — ir tai tik tokių grupių, kurios žmonių socialiniame, kultūros ar ūkio gyvenime turi didesnės reikšmės. Jei kurio žmogaus verslas ar kitas pažymys yra tiek savaimingas, kad nesudaro visuomeninės reikšmės grupę, tat paprastai toki vienetai priskiriami liekanoms, kurios sudaro „kitų“ ar „kitokių“, t. y. jokių pažymiu neaprėžtųjų, grupę. Dėl to, pav., nė vienos valstybės verslo statistikoj nėra atskiros grupės, pavadintos „ministeriai“, — nors jie ir galingi, ir kiekvienoj verslo statistikoj yra grupės, pavadintos „darbininkai“, „amatininkai“, arba net siuvėjai, kalviai ir t. t.

Apskritai imant, statistika neturi tikslo nei reikalo skaitmeninę fotografiją tiekti. Gyvenimas ir jo reiškiniai yra tiek įvairūs, kad visiškai pilnas ir visai tikras jo atvaizdavimas nėra galimas ne tik skaitmenimis, bet nei spalvomis, nei žodžiais, nei kitomis žmogui prieinamomis priemonėmis. Bet toks tobulas gyvenimo ir jo reiškinių atvaizdavimas niekam nėra nė reikalingas. Pavyzdžiui, jei pas mus buvo daromas visuotinis gyventojų surašymas, tai visai ne tam, kad sužinoti, kur gyvena, kada gimęs ar vedęs, ką ir kur dirba ir t. t. Petras,

Jonas ir Baltrus, bet tam, kad reikale būtų galima žinoti, kiek žmonių toj ar kitoj vietoj gyvena, kokia tų žmonių sudėtis ar socialinė padėtis, kaip jie tiki ar kalba ir t. t. Tas pats ir visur kitur, ir tai ne vien žmonių gyvenime. Atskiro žmogaus padėtis, išsimokslinimas, jo sielos pergyvenimai kartais gali būti svarbūs ir įdomūs jo artimiesiems, o jei žmogus prasikišęs ar visuomenės gyvenime svarbias pareigas eina — net ir tolimiems žmonėms arba visuomenei. Bet statistika ir tokiais atsitikimais neturi ko veikti, jei tų ypatybių niekur kitur nerandama arba jeigu jos nesudaro masės charakterio ir nesiduoda lyginamos. Tik charakteringosios, masės pobūdį nušviečiamos žymės ar jų pažymiai statistikos žymėtini ir tik tokios žymės statistikai reikšmės teturi; visa kita gali būti svarbu ir įdomu psichologui arba šiaip kuriam specialistui ar individualybių mėgėjui, bet ne statistikui.

Šiuo atžvilgiu statistiko darbas kai kurių lyginamas ir tikrai lygintinas su gamtininko sistematinimu ir menininko stilizavimu. Savo dalyką žinąs botanikas kiekvieną, kad ir dar taip savišką, žolelę priskirs kuriai nors augalų rūšiai, — nors nepatyrus akis gali nematyt panašumo. Geras dailininkas, piešdamas paveikslą, nežiūrės smulkmenų, o ims ir rodys tik charakteringuosius bruožus — arba tokius, kurių reikšmę jis ypač pabrėžti norės. Geras režisierius apleis visas menkos vertės veikalo ar veikiamo asmens charakterio smulkmenas, kad juo geriau pabrėžtų pagrindinę veikalo idėją ar duotų charakteringąjį tipą. Taip ir statistikas. Iš turimų duomenų ir pažymių jisai parinks tik pačius charakteringiausius ir turimą masę suskaidys arba sugrupuos taip, kad vienu akies apmetimu visą rūpimą ir tyrinėjamą dalyką, arba reiškinį, būtų galima visame platume pamatyti.

4. Statistikos mokslas ir gyvenimas.

Ta aplinkybė, kad statistikos metodai įgalina masės charakteringas ypatybes ir jų santykį, arba lyginamąjį svorį, pažinti, statistiką padaro pirmaeilės reikšmės disciplina tiek mokslui, tiek gyvenimui. Nėra ir negali būti pasaulyj žmogaus, kuris sugebėtų visus gyvenimo atsitikimus, ar šiaip reiškinius,

suskaityti arba suskaitytus galvoj palaikyti. O ir vidutinių gabumų žmogui visai nėra sunku susivokti, jei tie reiškiniai arba atsitikimai bus „stilizuoti", t. y. atitinkamai sugrupuoti ir parodytos tik pačios charakteringosios žymės. Pavyzdžiui imkime nors 1923 m. visuotiniojo gyventojų surašymo duomenis: 2.028.971 asmens lapelių, kur pažymėta kiekvieno atskiro žmogaus amžius, lytis, tautybė, padėtis šeimoj ir versle ir t. t., savaime sudaro tik popierio krūvą. Tačiau tuose lapeliuose sužymėtieji duomenys, suskirstyti atskiromis grupėmis arba tautybės, tikybos, amžiaus ir kitais pažymiais, lengvai gali būti surašyti ant vieno puslapio; ir čia mes galime matyti ne tik visos Lietuvos gyventojų sudėtį ir padėtį, bet net ir pasiskirstymą gyvenamomis vietomis. Dar aiškesnis bus vaizdas, jei šalia visų bendrųjų skaičių turėsime ir santykinius, pav.: kiekvienam 1000 vyrų išeina 1006 moterys; iš visų 2.028.971 gyventojų yra 83,9% lietuviai, 7,6% žydai, 3,2% lenkai ir t. t. Čia mes jau turime visai aiškų, tam tikrais pažymiais stilizuotą, gyventojų vaizdą, kuris gana lengvai duodasi ir galvoj palaikomas.

Bet čia vėl gali kas nors prikišti: o kam visa tai reikalinga? Kokia nauda iš visų statistikos duomenų ar skaičiais „stilizuotojo" vaizdo? Juk visa statistika dar nelabai senas dalykas, o žmonės vis dėlto gyveno. Štai filozofas J a m e s pasakoja, kad vienas anglų keliauninkas, susidraugavęs Turkijoj su vieno miesto k a d i, paprašė jo statistikos duomenų iš savojo miesto gyvenimo. Ir tasai kadi šitaip atsako:

„Mano kilnusis drauge! Tai, ko tu iš manęs reikalauji, yra tiek pat sunku, kaip ir nenaudinga. Nors aš visą gyvenimą šitoj vietoj praleidau, bet niekuomet man į galvą neateidavo teirautis nei kiek čia yra namų, nei kiek gyventojų; o apie tai, kiek kurie žmonės ant savo asilų prekių pakrauja, ar kiek jų į laivus sukrauja, — man visiškai nerūpi... Aš garbinu Dievą už tai, kad nieku nesiinteresuoju, kas man nėra būtinai reikalinga... O mano dūšia! O mano avinėli! Netyrinėk daiktų, kuriais tu nesuinteresuotas (kurie tau nenaudingi). Tu atėjai pas mus — mes tave širdingai priėmėm; ir keliauk sau sveikas[2]".

---

[2] Cituota iš *Schwartz*, Philosophie der Statistik (Allg. Statist. Archiv, 21 Bd., II. Heft, pusl. 206—7, 1931).

Nors ir gilios mintys atrodo šito turkų išmintingojo k a d i, vis dėlto labai galima abejoti, kad anas Anglijos keliauninkas būtų buvęs jo įtikintas ar atsisakęs statistikos duomenų pasiklausti apie lankomų vietų gyvenimą. Teisybė, yra nemaža žmonių dar ir šiandien ne vien tik Turkijoj, kurie niekuomet nėra pasiinteresavę, kiek jų kaime ar mieste yra namų ar gyventojų. Teisybė ir tai, kad toki žmonės puikiausia be tokių žinių apsieina. Tačiau į tai galima atsakyti, kad ir namai, vežimai ar daug kas yra ne nuo amžių ir kad žmonės ilgus laikus be jų apsieidavo, — vis dėlto niekam į galvą neateina ginčyti, kad tai ne naudingi dalykai.

Bet net jeigu atmesti visą tą pasitenkinimą, kurį teikia patys vieni gyvenimo reiškinių statistikos tyrinėjimai, tai statistika duoda dar ir tą naudą, kad iš jos duomenų galima patirti ne tik praeitį ir dabartį, bet dažnai spėti ir apie ateitį. Pavyzdžiui imkime nors ir mirimų statistiką. Niekas negali pasakyti, kada mirs bet kuris žmogus. Bet mes drąsiai galime sakyti, kad, pav., 1938—40 m. Lietuvoj numirs kas metai apie 30.000 —40.000 žmonių. Tas pats su gimimais, vedybomis ir t. t. Iš kur toksai įsitikinimas? Iš ten pat, iš kur mes esame įsitikinę, kad kas dieną iš ryto tam tikrą valandą ir minutę užtekės saulutė arba kad gruodžio mėn. 21 d. kas metai pas mus diena bus trumpiausia, o ne ilgiausia. Nes per metų metus buvo pastebėta, kad saulės tekėjimai, nusileidimai, dienos ilgumas ir nakties trumpumas nors ir keičiasi, bet, metams praėjus, vėl tie patys kitimai pasikartoja. Tas pats ir su gimimais, mirimais ir t. t. Nors čia tokio griežtai taisyklingo pasikartojimo nėra ir būti negali, — nes dienos ilgumas, saulės tekėjimai ir panašūs gamtos reiškiniai pareina nuo aiškiai aprėžtų ir (bent žmogaus akyse) nekintančių gamtos dėsnių. Tačiau gimimai, mirimai ar kt. panašūs reiškiniai priklauso veiksnių, kurie, žmogaus kontrolei nepasiduodami, vis dėlto yra tiek aprėžti, kad per daug didelių nukrypimų sudaryti neįstengia. Ir jau, pav., nuo 1921 m. šių skaičių (30—40 tūkst.) ribose mirimai Lietuvoje pasikartoja.

Čia dar vienas ypatingas dalykas reikia prisiminti. Senai pastebėta, kad, bendrai imant, berniukų gema truputį daugiau negu mergaičių (apie 21 bern. ir 20 merg.). Bet „jei imsime at-

skirus namus (šeimas) skaičiuoti, tai viename rasime gimus tiktai mergaites, kitame vien berniukus, — arba jei ir ne vienus, tai jau labai nelygioj proporcijoj. Nedaug terasime tvarkos ir mažose bendruomenėse arba atskiruose sodžiuose: tai, pavyzdžiui, miršta vienais metais du ar trys, tai vėl 6, o paskui gal 12 ar net daugiau. Kur čia tvarka ar dėsningumas?“[3] Vienok, jeigu paimti visą valsčių, apskritį ir ypač valstybę, tai pamatysime tokį per metų metus pasikartojantį dėsningumą, kurį tas pats autorius (Suessmilch) ne be reikalo kareivių pulko rikiuotės žygiui prilygino.

5. Statistikos dėsningumai.

Čia mes jau susiduriam su vad. statistikos dėsningumais bei pareinamybėmis. Tiek gamtoje, tiek žmonių visuomenės gyvenime niekas neatsitinka be priežasties, o viskas pareina nuo anksčiau buvusių ir esamų veiksnių įtakos. Žemė sukasi apie saulę, mėnulis — apie žemę; taip pat kitos planetos ir žvaigždės savo keliais eina ne atsitiktinai, bet tam tikriems fizikos dėsniams pasiduodamos. Tuos dėsnius pažindami mes galime iš anksto atspėti dangaus kūnų kelius ir — saulės bei mėnulio užtemimus. Oro atmainos mūsų žemėje taip pat eina dėsningai. Ir nors iš anksto ilgesniam laikui oro atmainų visai tiksliai atspėti negalima, vis dėlto niekas iš mūsų nelaukia ir nesitiki liepos mėnesį sulaukti sniego ar šalčio, arba sausio mėnesį karščių, — nežiūrint, kad metraščiai ir rodo senovės Lietuvoj kai kada vasarą buvus vėsaus oro ir net sniego, o žiemą — šilimos. Visai tas pats apsireiškimas ir žmonių visuomenės gyvenime. „Moraliniai ir dvasiniai apsireiškimai, jei žiūrėt visos masės, taikosi gamtos apsireiškimams, ir... juo yra didesnis tyrinėjimų (observuojamų) vienetų (individų) skaičius, tuo silpniau matomos individualinės fizinio arba moralinio pobūdžio ypatybės ir tuo daugiau reiškiasi bendrieji veiksniai, kuriais remiasi visuomenės būtis ir jos gyvata. Tik kai kuriems ypatingai genialiems asmenims yra lemta padaryti apčiuopiamos

[3] *Suessmilch,* Göttliche Ordnung... Berlin, 1765. I, 56.

įtakos visuomenei, o ir ta įtaka yra dar reikalinga laiko, kad pilnai pasireikštų"[4].

Tai yra žodžiai paties moderniosios statistikos įkūrėjo Quetelet, — ir jie tolimesnio aiškinimo čia nereikalingi. Toliau mes pamatysime, kad šitas Quetelet žodinis apibrėžimas atitinka ir matematiškai pagrindžiamą vad. didžiųjų skaičių dėsnį iš tikimybių teorijos. O tai yra svarbi aplinkybė. Nes jeigu pakankamai didelis tyrinėjamų vienetų skaičius geriau ir aiškiau pabrėžia bet kurios masės charakteringąsias ypatybes, o užstelbia necharakteringas, arba antraeiles, tad tuo pačiu duodama į rankas ir matematiškai pagrindžiama bei aprėžiama matavimo priemonė masėms, arba masiniams reiškiniams, tyrinėti.

Trumpais žodžiais statistikos esmę galima šitaip apibrėžti:

1. Statistika nėra paprastas atsitikimų, reiškinių ar kokių daiktų arba dalykų skaičiavimas; statistika yra masių arba masinių reiškinių žymėjimas skaitmenimis ir jų saistymas (klasifikavimas) arba stilizavimas charakteringomis žymėmis bei pažymiais (ypatybėmis).

2. Statistikai pagrindą duoda vadinamas didžiųjų skaičių dėsnis; nes tik ganėtinai dideliam tyrinėjamų atsitikimų arba reiškinių skaičiui esant, charakteringosios žymės (ypatybės) pilnumoj pasireiškia.

3. Didžiųjų skaičių dėsningo veikimo charakteringųjų žymių ir pažymių aikštėn iškėlimas veda prie vadinamųjų statistinių dėsningumų — laiko, vietos, prigimtinių bei socialinių sąlygų santykiavimuose. Šitie statistiniai dėsningumai rodo įtaką tų veiksnių, kurie nulemia statistikos stebimųjų masių, arba masinių reiškinių, charakteringas ypatybes.

4. Statistiniai dėsningumai įgalina daryti tikimybių spėjimus; tik šiuo atveju tenka atsižvelgti ne tik pagrindinių ypatybių charakterio (tipo, normos), bet ir esamų (arba galimų) nuo normos (tipo) nukrypimų arba patikimybės paklaidų (patikimoji paklaida).

---

[4] *Quetelet,* Sur L'Homme et de Développement de ses Facultés. Paris, 1835. I, 12.

B. Statistikos surašinėjimų istorinė apžvalga.

6. Surašinėjimai ir statistika senovės gadynėj.

Statistikos esmę mes lengviau pažinsime iš statistikos istorijos bei jos pavyzdžių. Štai Kalėdų metu skaitomos Šv. Luko Evangelijos antrojo skyriaus pradžioje pasakyta:

„Anuo metu išėjo Ciesoriaus Augusto paliepimas surašyti visą pasaulį. Tas pirmas surašymas buvo padarytas Syrijos valdovo Kyrino. Ir visi ėjo įsirašydintų, kiekvienas į savo miestą. Ėjo taipogi ir Juozapas iš Galilėjos iš Nazareto miesto į Judėją į Dovydo miestą, kurs vadinamas Betliejus, nes jis buvo iš Dovydo namų ir giminės, kad įsirašydintų draug su pažadėta sau moterimi Marija, kuri buvo neščia“[5].

Iš čia matome, kad Kristaus gimimo metais Romos imperijoj (Romos imperatoriai tada laikė save viso pasaulio valdovais, dėl to ir Evangelijos žodžiai kalba apie viso pasaulio surašymą) jau buvo daromas visuotinis gyventojų surašymas. Ir, kaip matyti, tasai surašymas buvo privalomas visiems Romos imperijos gyventojams, kurie net turėjo keliauti į savo gimimo vietą, „kad įsirašydintų“. Klaidinantis bet gi yra Šv. Luko Evangelijos antrasis sakinys, kuris šitą surašymą vadina pirmuoju surašymu. Nes ciesoriaus Augusto paliepimu darytasis visuotinis gyventojų surašymas ne tik neapėmė viso pasaulio, bet ir Romos imperijoj buvo toli gražu ne pirmutinis. Daugiausia — šitas surašymas galėjo būti pirmutinis Sirijoj ar Judėjoj, arba kitose Romos imperijos provincijose. O pirmasis Romos gyventojų surašymas siejamas su Serviaus Tuliaus vardu. Jau 443 metais prieš Kristaus gimimą buvo įsteigtas vadinamas cenzorių institutas. Šitie „cenzoriai“, be kita ko, turėjo žiūrėti ir gyventojų „cenzo“, t. y. įrašytųjų arba surašytųjų stovio pažymių. Patys surašinėjimai Romos respublikos laikais buvo daromi kas penkeri metai. Tam tikslui visi Romos piliečiai, nors ir toliausia gyventų, turėjo į Romą atvykti. Šis gimtosios vietos surašomasis pažymys, matyti, vėliau buvo ir visos imperijos gyventojų surašymams pritaikintas. Ir surašinėjami buvo ne vien patys gyventojai, bet ir jų

[5] Cituota iš *Skvirecko* „Evangelijos dalys“, 5.

turtas arba inventorius. Vergai tais laikais į gyventojų surašymo duomenis nepatekdavo — jie patekdavo į inventoriaus sąrašus.

Pirmieji surašymai būdavo daromi tiktai pačioj Romoj, bet vėliau jau ir kai kuriose provincijose. Imperatorių laikais surašymai buvo išplėsti visoj imperijoj. Bet dabar jie jau rečiau būdavo daromi: kas dešimtis metų (anksčiau kas penketas metų). Laikui bėgant surašymai buvo daromi dar rečiau ir nereguliariai; galų gale jie išnyko visai. Iš to, kas yra žinoma apie Romos gyventojų surašomuosius pažymius, galima teigti, kad tie surašinėjimai būdavo daromi valstybinės administracijos ir ypač mokesčių bei prievolių reikalui.

Tą patį galima sakyti ir apie buvusius ž y d ų gyventojų surašinėjimus. Mozės IV knyga kaip tik prasideda instrukcijomis visuotiniam gyventojų surašymui daryti. Surašymas buvo daromas Sinajuj. Surašinėjami buvo visi vyriškiai — pirmoj eilėj tie, kurie turėjo 20 metų ir tiko karo tarnybai (IV, Mozės, I, 3). Tačiau buvo surašinėjami ir visi kiti vyriškiai, pradedant vieno mėnesio kūdikiais (IV, Mozės, III, 15). Surašoma buvo kilmės ir giminystės pažymiais.

Mozės darytieji žydų tautos gyventojų surašymai taip pat buvo daromi administracijos reikalui. Tų surašymų duomenimis pasirėmus buvo sudarytas valdžios aparatas, uždedamos prievolės, renkami mokesčiai bei rinkliavos („Imk po penkis saikus nuo galvos... ir duok pinigus Aronui ir jojo sūnams“ ir t. t. IV, Mozės, III, 47—48).

S e n o v ė s G r a i k i j o j reguliarinių visuotiniųjų gyventojų surašinėjimų Romos pavyzdžiu, rodos, nėra buvę. Bet laiks nuo laiko būdavę nereguliariniai gyventojų ir jų turto surašymai mokesčių apdėjimo reikalui. Be to, senovės Graikijos miestuose buvo vedama gyventojų registracija (rejestrai).

E g i p t e pirmasis visuotinis gyventojų surašymas buvęs apie 3.050 m. prieš Kristaus gimimą. Aplamai, Egipto gyventojų surašinėjimai siejami su viešaisiais darbais arba tų darbų sumanymais (piramidžių statyba, irrigacijos darbai ir kt.). Vad. Romos gadynėj būdavę daromi kas 14 metų visuotinieji gyventojų surašymai mokesčių (pagalvės ir kt.) rinkimo reikalui. Be to, Egipte iš senovės buvo pažįstamas policinis gy-

ventojų prisimeldavimas, kuris taip pat teikdavo medžiagos gyventojų statistikai.

K y n u o s e pirmieji gyventojų surašymai statistikos istorikų parodymais turėtų būti buvę maždaug 2.300 m. prieš Kristaus gimimą, o 550 m. prieš Kr., šalia gyventojų surašymo, buvę daromi ir žemės plotų matavimai, taip pat žemės valdymo ir pramonės statistika. Bet Kynų žemės ūkio istorija rodo, kad toki surašymai galėjo ar net turėjo būti daromi ir anksčiau. Taip bent duoda pagrindo manyti Kynų ūkio istorijoj žinoma žemės valdymo ir naudojimo Tsing Tien sistema, kurios pagrindan buvo padėta matematiniu tikslumu griežtai aprėžti žemės sklypai, tam tikroj tvarkoj sugrupuoti ir tam tikram žmonių skaičiui paskirti. Nežinant gyventojų skaičiaus nė žemės plotų didumo, tokiais pagrindais žemės valdymo ir naudojimo reformos įgyvendinimas atrodo neįmanomas. O pati Tsing Tien sistemos pradžia siekia 2.737 metų prieš Kristaus gimimą[6].

I n d i j o j, rodos, senovėje buvo daromi daugiausia turtų arba ūkio, o ne gyventojų, surašinėjimai.

J a p o n i j o j pirmieji gyventojų surašymai buvę daromi prieš porą tūkstančių metų.

Aplamai atrodo, kad gyventojų surašymai ir, apskritai, statistikos praktika senovės tautoms ir valstybėms buvo gana gerai pažįstama. Ir buvo pažįstami ne vien atsitiktiniai ar neperiodiniai surašinėjimai, bet ir gerai organizuoti, tam tikrų statistikos organų atliekami ir viešosios teisės organų gyvendinami visuotinieji gyventojų ir jų turto surašymai — ar net žemės plotų matavimai su skirtingų rūšių ir kt. pažymėjimais. Šitai juo labiau įsidėmėtina, kad šių dienų gyventojų ir kt. statistika visai dar nesena, — ir tai ne tik naujosiose Europos valstybėse, bet ir senovės rytų kultūros kraštuose, kur statistikos praktika gilioje senovėje buvo gana tobulai pastatyta.

Apskritai imant, apie Kristaus gimimo laikotarpį iš senovės pažįstami gyventojų ir kt. surašinėjimai ir iš viso statistikos praktika (statistikos mokslo arba teorijos senovėj, rodos, ir būti nebuvo) išnyksta. Tik viduramžiais kai kuriose šalyse pasitaiko vieno antro miesto visuotinis gyventojų surašymas,

---

[6] *Mabel Ping-Hua Lee,* The Economic History of China, N. Y. 1931 p. p. 33—38.

arba vienas kitas feodalas mėgina padaryti savo turtų arba inventoriaus sąrašus. Taip pat atsirasdavo vienas ar kitas vienuolis, kuris atsitikusiam reikalui, arba net savo smalsumui patenkinti, suskaičiuodavo savo pažįstamų vietų vardus arba šiaip kokius daiktus.

7. Viduramžių surašinėjimai.

Visur, kur iš senovės randame organizuotos statistikos pėdsakų, toji statistika sutampa su valstybinės valdžios įsigalėjimu. Ir statistikos darbai daromi valstybinės centro valdžios įsakymais ir jos reikalams pritaikinti (karo arba apsaugos reikalui, mokesčiams arba prievolėms atlikti, administracijos arba socialinių reformų reikalui).

Visai kitoks vaizdas Kristui jau gimus. Senovės valstybinės organizacijos arba iširo, arba visai nusilpo. Visuotinieji sumišimai ir klajokliniai tautų žygiai bei naujų vietų ieškojimai, taip pat religiniai perversmai ir juos lydį karai, visą gyvenimą darė tiek nepastovų, kad koki nors gyventojų ar kt. statistikos surašinėjimai buvo negalimi ir nereikalingi. Net vėliau, visam gyvenimui šiek tiek aprimus ir susikūrus feodalinėms valstybėms, sistemingi statistikos darbai nebuvo būtini. Nes feodalas kunigaikštis savo vasalus pažino ir be statistikos. Taip pat vasalai pažino savuosius ricierius ar bajorus su baronais, kurie, vergų arba tarnų padedami, ir visos nepainios natūrinio ūkio tvarkos žiūrėjo. Taigi, nenuostabu, kad pirmasis ir, galima sakyti, vienintelis viduramžių statistikos darbas buvo atliktas toj šalyj, kuri centralizuotos valstybinės organizacijos atžvilgiu tais laikais pirmojoj eilėj stovėjo (ar bent pastoviausiose sąlygose buvo), t. y. Anglijoj. Vad. Vilhelmo Nugalėtojo laikais 1080—1087 m. tenai buvo daromi surašinėjimai, ir visa kas įtraukiama į tam tikrą knygą, kuri šiandien vadinama Domesday Book. Tai nėra koks statistikos žinių metraštis nei visuotinio surašymo žinių rinkinys ta prasme, kaip dabar yra priimta panašių žinių rinkinius leisti. Šita knyga išsyk, rodos, neturėjo net jokio pavadinimo. Dabar žinomas „Domesday Book“ pavadinimas, jei Henry Ellis aiškinimas

yra teisingas, pareina nuo to, kad kalbamoji knyga buvo padėta ir saugojama karaliaus iždinėj, o pati iždinė tais laikais buvusi Wincesterio ar Westminsterio bažnyčioje, atseit, Dievo namuose — Domus Dei[7].

Šalia Domesday Book, ta pati knyga turi dar apie aštuonis pavadinimus (iš jų plačiausia žinomas „Liber censualis Wilhelmi Primi Regis Angliae), kurie bus bene visi pridėti jau vėliau įvairių knygos naudotojų.

Savo turiniu ir statistikos duomenų tvarka Domesday Book labai primena 1528 m. „popis zemsky" Lietuvoje. Surašyta tenai visos atskiriems valdytojams priklaususios vietovės su pažymėjimais, kiek kurios rūšies yra žmonių, kiek jie turi žemės ir kokios toji žemė rūšies, kokios yra prievolės, koki ir kaip dideli pajamų šaltiniai ir t. t.

XIV—XVI š. įvairiuose Europos kraštuose darytieji dvarų inventoriai ar vad. breviarium fiscalium, missi, Karoliaus Didžiojo karių sąrašai, taip pat Brandenburgo, Neumarko ir kt. surašinėjimai nei turiniu, nei statistiniu vertingumu negali prilygti 1080—1087 m. Domesday Book. Iš Vakarų Europos viduramžio sąrašų ar surašymų didesnės statistinės vertės turi tik 1449 m. Nuernbergo miesto gyventojų surašymas, 1501 ar 1507 m. Sicilijos sąrašai ir 1526 ar 1527 m. Romos gyventojų surašymas.

Apie pusę XVI š. buvo surašinėjama ir kai kurių kitų Italijos miestų gyventojai (1540 m. Venecijoj, 1559 m. Florencijoj ir k.). 1571 m. Saksonijos kurfiurstas Augustas išleido „taisykles ir pranešimą, kaip kunigaikštystė ir kraštas gali būti aprūpintas javais ir kitomis reikmenimis". Tenai, tarp kita ko, valdininkams ir teisėjams buvo uždėta pareiga paduoti „žinias ir skaičius" apie visus ūkius ir jų gyventojus — vyrus, moteris, vaikus ir t. t. Bet kiek šitas kunigaikščio įsakymas buvo pildomas, sunku pasakyti, — nors už nepildymą ir bausti žadėta (po 10 fl. už kiekvieną neįrašytą asmenį).

---

[7] „in a place called Domus Dei, or God's House, and so the name of the booke, therefore called Domus Dei, and since, shortly, Domesday *(Henry Ellis,* A. General Introduction to the Domesday Book, London, 1833. I. 47).

8. Lietuvos surašinėjimai ir 1528 m. „popis zemsky".

Statistikos surašinėjimo darbų būta ir D. Kunigaikščių Lietuvoje. Neskaitant įvairių daugiau ar mažiau vietinės reikšmės, be bendro plano, darytų gyventojų „tarnybų" arba jų turto surašinėjimų, kurie buvo praktikuojami nuo Vytauto laikų, 1507 m. buvo nutartas, bet tik 1528 m. įgyvendintas visuotinis Lietuvos dvarų bei sodžiaus gyventojų turto bei prievolių surašymas. Šito surašymo duomenimis remiantis paskui buvo padaryti Lietuvos Metrikoj išlikę Lietuvos karo pajėgų sąrašai. Tais duomenimis remiantis taip pat buvo naujai pertvarkyta ir visa eilė kitų prievolių, mokesčių ir net padaryta administracinės technikos reformų.

Deja, paties surašymo originaliniai duomenys yra išlikę tik kai kurioms Lietuvos vietoms. Bet iš duotosios instrukcijos surašinėtojams, taip pat iš užsilikusių atskirų vietų aprašymų, aiškiai matyti, kad surašymas buvo iš anksto gerai suplanuotas. Statistiškai metodologiniu atžvilgiu Lietuvos 1528 m. „Popis ziemsky" net geresnis už Anglijos Domesday Book įrašus. Nes čia buvo surašyti ne tik gyventojai, jų padėties ar tarnybos pažymiais, bet daugely atsitikimų ir moterys, vaikai, taip pat — ir tai taisyklė — jų laikomų gyvulių skaičius šalia atskiro pažymėjimo, kiek yra „jungų" arba „žagrių", kiek žmonių žemę dirba, kiek mokesčių moka ir kokios rūšies prievolių turi (lažas, čyžė ar kt.), taip pat kokius javus augina ir t. t. Tokios rūšies žinių Anglijos Domesday Book man užeiti neteko — ten pasitenkinama žemės naudojimo pažymiais (terra, silva ar nemus, pastura, pratum ir p.), o apie gyventojus tik pasakyta, kad tiek ir tiek esą villani, censarii arba servi, cortarii, bordarii — be pažymėjimo apie šemynos narių skaičių ir be inventoriaus.

Šiek tiek modifikuotas, bet 1528 m. panašus, surašymas dar pakartotas 1565 ar 1567 metais. Be to, 1557 m. ryšyj su valakine žemės reforma įvestosios vad. „ličbos", t. y. metinės ūkio apyskaitos, savaime duoda daug svarbios statistinės medžiagos ir Lietuvos ūkiui nušviesti. O Lietuvos Metrikoj išlikę Lietuvos kariuomenės sąrašai, paremti 1528 m. visuotiniojo

surašymo duomenimis, ir dabar dar savo pilnumu ir platumu yra vienintelis statistikos žinių šaltinis Lietuvos anų laikų karo jėgoms pažinti[8].

9. Naujųjų laikų statistika.

Aplamai, statistika mėgsta pastovią ir gana tvirtą valstybinę organizaciją. Taip buvo senovėje, taip pasirodė viduramžiais ir pasitvirtino ypač vad. naujaisiais laikais. Viduramžiais anksčiausia susitvarkė Anglija — ji patiekė ir seniausią tobulesnės statistikos dokumentą, vad. Domesday Book. Lietuvoje 1528 metų surašymas ir jojo visas turinys taip pat ne atsitiktinis. Nuo Vytauto gadynės Lietuva buvo tiems laikams gana stipriai organizuota valstybė. Taigi visai suprantama, kad buvusioj padėtyj ir pribrendus socialinėms bei fiskalinėms reformoms, taip pat dėl bendrosios politikos, atsirado reikalas padaryti visuotinį gyventojų turto arba ekonominio pajėgumo surašymą. Vakarų Europos didesniuose plotuose pastovesnis valstybinis gyvenimas iš tikrųjų prasidėjo tik vadinamiems naujiesiems laikams ateinant su visais jų naujai keliamais klausimais bei reikalais. Ne tik valstybių merkantilistinei politikai buvo reikalinga turėti tikslesnių statistikos duomenų apie valstybės ekonominę padėtį. Religiniai sambrūzdžiai ir ypač reformacijos iššauktos religinės kovos privedė prie reikalo registruoti savosios tikybos išpažintojus. Triento koncilija (1543—1563) įveda visur privalomą krikšto, jungtuvių ir laidotuvių registraciją ir kunigams uždeda pareigą daryti visuotinius tikinčiųjų surašymus.

Pasaulinės valdžios organai tais laikais perdidelio susidomėjimo vien gyventojų statistika nerodė. O jei ir būdavo daromi gyventojų surašinėjimai, tai ūkiais („dūmais“) arba šeimynomis. Bendrasis gyventojų skaičius būdavo tik spėjamas (ūkių arba dūmų skaičius dauginama 4 ar 5 ir tariama, kad tai turi atitikti tikrąjį gyventojų skaičių). Išimtį daro tiktai Sicilija, kur 1623, 1651, ir 1681 metais buvo daromi visuotini gy-

[8] Žiūr. *Liubavskij*, Oblastnoje delenije i miestn. upravl. Litovsko-Russk. Gosud., priedų Nr. 14, *Rimka*, „Lietuvos visuomenės ūkio bruožai“, psl. 46, *A. Tarulis*, Didž. Lietuvos kunigaikštijos karo prievolių (1528 m.) sąrašų patikimumas statistikai („Ekonomika“, 1936, 2, psl. 122 ir sek.).

ventojų surašymai, kurie, tačiau, dar toli gražu neatitinka statistikos mokslo reikalavimų. Statistinės metodologijos reikalavimus atitinkąs visuotinis gyventojų surašymas buvo pradėtas ne Europoj, bet Amerikos Jungtinėse Valstybėse. Ten 1788 m. konstitucijos aktu nustatyta, kad visuotinieji gyventojų surašymai būtų daromi kas 10 metų natūriniais ir socialiniais gyventojų sudėties pažymiais. Pirmasis toksai „census" Amerikos Jungtinėse Valstybėse buvo padarytas 1790 metais ir reguliariai kas 10 metų kartojamas ligi pat naujausiųjų laikų.

Europoj panašiais pagrindais padarytieji pirmieji visuotini gyventojų surašymai buvo 1801 metais Prancūzijoj ir Anglijoj. Berods, pradedant 1710 metais Prūsuose buvo žinomi vadinami gyventojų lakštai („Populationslisten", vėliau „Historische Tabellen"), kur taip pat būdavo paduodami ir gyventojų skaičiai. Bet tai buvo daugiau anketinės žinios, o ne visuotiniojo surašymo duomenimis paremta gyventojų statistika. Be to, tos žinios būdavo renkamos įvairiais keliais ir metodais, ir tuo būdu paduotieji duomenys sunkiai davėsi lyginami. Gyventojų statistikai daug tinkamesnis surašymas Vokietijoj buvo daromas XIX a. trisdešimtaisiais metais muitų sąjungos reikalui (muitų pajamoms paskirstyti). Visuotiniojo gyventojų surašymo keliu gautoji gyventojų statistika Vokietijoj tenka 1871 metams drauge su Vokiečių Jungtinės Valstybės (Reicho) pradžia.

L i e t u v o j pirmasis visuotinis ir tiesioginis gyventojų surašymas buvo 1897 metais drauge su visos Rusijos pirmuoju visuotiniu surašymu (Mažojoj Lietuvoj toki surašymai būdavo kas 5 metai drauge su Vokietijos gyventojų surašymais). Antrasis Lietuvos gyventojų surašymas buvo 1923 metais, o Klaipėdos krašto 1925 m. Žinios apie Lietuvos gyventojų skaičių prieš pirmąjį surašymą, taip pat ir vėliau Rusų metraščiuose ar kt. paduotieji duomenys, buvo paremti netiesioginiais skaičiavimais arba vertinimais.

Pirmasis visuotinis žemės ūkio surašymas Lietuvoje padarytas 1930 metais.

C. Statistikos mokslo istorinė apžvalga.

10. Statistika senovėje ir viduramžiais.

Jau buvo sakyta, kad statistikos pradžia ir kilimas yra susiję su valstybinės organizacijos pradžia ir kilimu. Taip pat matėme, kad statistikos surašinėjimai daugiau buvo paplitę kraštuose, kur buvo susikūrusios senovės kultūros valstybinės organizacijos, negu vėlesniais laikais, kai tos senosios valstybinės organizacijos išklero ar buvo išardytos, o joms atitinkamų naujų dar nebuvo sukurta. Bet vaizdas pasikeičia, jei imti ne statistikos praktiką, t. y. statistikos žinių rinkimą, o statistikos mokslą. Čia senovė, statistinės praktikos srity davusi nemaža pavyzdžių, neteikia mums beveik nieko. Berods, šen ar ten vienas ar antras veikėjas valdžios įstaigų surinktą statistikos medžiagą naudoja savo raštams arba aprašymams. Šioj srity yra žinomi darbai tūlų senovės graikų, kurie aprašė kai kurių tautų institucijas arba kai kurių miestų (Spartos, Atėnų, Korinto ir kt.) gyvenimą. Patsai Aristotelis buvo aprašęs net 158 miestų valstybes, o Romos Augustas yra laikomas autorium pagarsėjusio Breviarium Augusti, kuris buvęs tęsiamas ir jo palikuonių. Visuose tuose rašiniuose buvo apsčiai ir statistinės medžiagos, paremtos administracijos archyvų ar net tiesioginių surašinėjimų duomenimis. Bet ligi šiol, rodos, nesurasta iš senovės gadynės nė vieno veikalo ar šiaip dokumento, kuris rodytų, kad to laiko mokslo žmonės būtų labai mėgę po ranka turimą statistikos medžiagą analizuoti ar mokslo metodais kokių nors statistinių dėsningumų ieškoti. Nėra net statistinių duomenų kritikos, nekalbant jau apie statistikos metodologiją. Ir taip buvo ne tik senovėje, bet net ir ankstybojo viduramžio laikais.

11. Renessansas ir vad. universitetinės statistikos kilimas.

Šiokią tokią atmainą padarė tik bebrėkštanti vad. Atgimimo arba Renessanso gadynė. Senovės kultūros tėvynėje imta ieškoti gyvenimo galios ir garbinti visa, kas tik ten buvo ar, atrodė, kad būta. Georg von Mayr sako, kad šitas palinkimas (kitur pavyzdžių ieškoti) kaip tik ankštai susi-

jęs su išsiplėtimu vadinamos vokiečių universitetinės statistikos su jos vad. „valstybiniais ypatumais“ arba „Staatsmerkwuerdingkeiten“ (notitia rerum publicarum). Pasak Mayr ir kt., dar 1420 m. Mocenigo patiekia memorialą apie Venecijos prekybos santykius. 1542 m. Contarini paskelbia net penkias knygas apie „De magistratibus et republica Venetorum“. 1562 m. Sansovino išleidžia rinkinį apie 20 senovės ir naujųjų valstybių su statistiniais daviniais. 1589 m. Botero skelbia savo „Relazioni universali divisi in quattro parti“ su teritorijos, santvarkos ir tikybos pažymių nušvietimu. Panašioj kryptyj dirbo ir popiežius Pius II (Aeneas Silvius) ir kt. 1544 metais išeina Sebastiano Muensterio Kosmografija. XVII š. pradžioj Prancūzijoj Pierre d'Avity išleidžia „Les états empires et principautés du monde“. Prie tos pačios rūšies priklauso ir Amsterdame brolių Elzevirų metraščių formos leistosios ir Jan de Laet redaguotosios respublikos („Respublicae Elzevirianae“).

Į visus šituos be jokios tvarkos ir sistemos skelbiamus statistikos duomenis, arba, teisingiau, užrašus šiek tiek naujos dvasios įnešė vadinama „vokiečių universitetinė statistika“ 17—18 šimtmetyj. Čia reikia prisiminti ir bendrąją to laiko dvasią bent Vakarų Europoje. Tai buvo vad. feodalizmo gadynė. Visur viešpatavo gana stiprūs feodalai (išskiriant Lenkiją ir Lietuvą, kur valdovas buvo bajorų statytinis, ir Angliją, kur įtakos įgavo lordų parlamentas). Valdovų reikalavimai, pageidavimai, o ir patys reikalai, didėjo ir plėtėsi. Ne tik patsai valdovas feodalas, bet ir jo valdininkai ir administracijos įstaigos buvo reikalingos visos eilės informacinio ir statistinio pobūdžio žinių. Tokių žinių laikui einant administracijos įstaigose prisirinkdavo ir be tiesioginių surašymų. Bet visos tos mažiau ar daugiau statistiškai informacinio pobūdžio žinios tuomet buvo laikomos lyg didžiausia valstybės paslaptis. Skelbti tuos duomenis viešai — tai reiškė beveik tiek pat, kaip išduoti savojo valdymo paslaptį kaimynui rivalui, o gal ir priešininkui.

Tuo tarpu valstybės įstaigų arba atskirų valdovų archyvuose prisirinkdavo ir tokių žinių bei duomenų, kuriais buvo galima, valstybės valdymo paslapčių visai neišduodant, ir pasididžiuoti. Taip pat, valstybės organizacijai besiplečiant, vals-

tybės organams dažnai buvo reikalinga po ranka turėti informacinio ir statistinio pobūdžio žinių, kurios jokios valdymo paslapties sudaryti negalėjo, arba, laikui bėgant, jau buvo pasikeitusios pažiūros dėl pačios „paslapties“ sąvokos.

Tokiose aplinkybėse visa eilė žmonių, ypač universitetų profesorių, atsiduria slaptuose valdovų archyvuose, ir netrukus ima rodytis mažesnės ar didesnės vertės veikalų su statistiniais įvairių įstaigų aprašymais. 1544 m. išėjusi Muensterio Kosmografija jau turi šios rūšies rašinių pažymį. Bet šiuo atžvilgiu charakteringas veikalas yra Seckendorfo „Teutscher Fuerstenstaat“. Tai jau gana sistemingas Vokiečių kunigaikštysčių ir jų institucijų aprašymas (taip pat pačių valdovų ir jų šeimos narių aprašymas). Dar toliau nuėjęs savo paskaitose prof. Conring (miręs 1681 m.), kurio „rerum publicarum“ buvo išleista iš klausytojų užrašų pačiam autoriui mirus. Halles prof. Schmeitzel savo paskaitų kursą (1747 m.) pavadino politine statistika. Vienok tikruoju universitetinės statistikos „tėvu“ yra laikomas Achenwall.

12. Achenwall'io „statistika“.

Goettingeno universiteto prof. Achenwall 1749 m. išleido veikalą vardu „Abriss der neusten Staatswissenschaft der vornehmsten Europaeischen Reiche und Republiken zum Gebrauch in seinen Vorlesungen“. Šiame veikale pirmą kartą „įpilietintas“ statistikos terminas, nors tais laikais žodis statistika naujiena nebebuvo. Patsai autorius sako, kad „vadinamos statistikos, t. y. atskirų valstybių mokslo sąvoka įvairiai aiškinama, ir didelėj raštų daugybėj nelengva rasti bent vieną aiškinimą, kuris savo sudėtinėmis dalimis bei konstrukcija atitiktų kitą aiškinimą“. Patsai autorius mano, kad valstybių mokslas arba statistika turėtų „tikrai parodyti piliečių visuomenės ypatingumus“ (Merkwuerdigkeiten), „ypatinga“ gi esą tai, kas geriau tinka valstybės gerovei“. Dėl to iš „begalinės“ tų ypatingumų daugybės turį būti parinkti tik toki, be kurių negalima pažinti tikrosios tautos gerovės. O viso to darbo tikslas — „daryti išvadas, kuo būdu valstybė gali būti išmintingai valdoma, t. y. kaip tas žinias pritaikinti politikai“.

Kaip matome, Achenwall'io statistika iš tikrųjų buvo visai ne statistika, kaip ją suprantame šiandien, bet greičiau taikomoji ekonomika arba ekonominė politika, o dar teisingiau — taikomoji politika. Berods, Achenwall lietė beveik visas sritis, kurias liečia ir šių dienų statistika. Bet, anot Mayr'o, šitas „statistikos tėvas" savo veikale jei ką įrodęs, tai stebėtiną statistikos skaičių vengimą. Viskas pas jį aprašyta nieko neapleidžiant, tarytum tai būtų bibliotekos knygų katalogas ar inventoriaus sąrašas; ir viskas nušviečiama utilitaristinės politikos atžvilgiu, — prie tam dar subjektyviškai, neatsižvelgiant į galimas ir net esamas kitas aplinkybes bei įvairius gyvenimo reikalavimus.

Achenwall į kiekvieną valstybę žiūrėjo kaip į vienetą su savais valstybiniais „ypatingumais". Dėl to jis ir aprašė kiekvieną kraštą atskirai monografiškai etnografiniu metodu. Bet jau B u e s c h i n g, kitas to paties Goettingeno universiteto profesorius ir tos pačios universitetinės statistikos atstovas, praskynė naują kelią: jis tą pačią medžiagą jau lyginamuoju metodu tvarkė ir klasifikavo žiūrėdamas ne tiek formališkai etnografinių, kiek esmės pažymių. Be to, Buesching daug daugiau dėmesio kreipė į skaitmeninius duomenis.

Visa eilė įvairių universiteto profesorių dirbo ta pačia kryptimi. Tik vieni sekė Aschenwall'į, kiti daugiau linko prie Buesching'o lyginamojo metodo. Šiek tiek įvairumo į visą tą kryptį įnešė S c h l o e z e r (1735—1809), kuris universitetinės statistikos metodus mėgino teoretiškai pagrįsti. Kad statistika yra tik bendrosios valstybinės politikos dalis, tai Schloezer taip pat neabejojo. Bet jam „valstybinių ypatingumų" moksle ir net pačioj jo sąvokoj nebuvo nieko absoliutaus, o viskas reliatyvu — priklausė aprašomos valstybės didumo. Valstybės arba jos atskirų kraštų aprašymas, istorija, sausi geografijos duomenys — visa tai esą nieko bendra neturi su statistika. Bet užtat politinė geografija sudaro statistikos pagrindą: nes ji rodanti valstybės gerovės laipsnį ir galią. „Statistika yra stovinti istorija, o istorija kintanti arba tęstinė (fortlaufende) statistika".

Šis Schloezer'io posakis vėliau gana dažnai buvo cituojamas, o ir dabar naudojamas statistikai charakterizuoti; nors, kaip matėme, jo autoriui statistika vaizdavosi šiek tiek kitaip.

Statistikos medžiagai tvarkyti Schloezer'is išstatė net rodyklinę tezę: pagrindinę galią sudaro žmonės, žemė, produktai ir pinigai; valstybės santvarka arba konstitucija šias jėgas sujungia; ir, pagaliau, valdžios organuose (vyriausybėj ir administracijoj) pasireiškia tos galios veikimas. Šito prasmingo rodiklio privalanti ir statistika.

13. Universitetinės statistikos pasekėjai ir jų ginčai.

Apie universitetinės statistikos žymesniuosius atstovus spietėsi visa eilė pasekėjų arba mokinių ne tik Vokietijoj, bet ir kituose kraštuose: Italijoj — Padovani, Tomassia Gioja, Romagnossi; Prancūzijoj — Peuchet, Donnaut, Ballois, Herbin; Anglijoj — Sinclair ir t. t. Jų tarpe, tačiau, būta nemažo nuomonių skirtumo dėl metodų, sąvokų ir kt. Būta taip pat ir jų metodo priešininkų bei kritikų. Griežčiausiu iš jų laikomas tūlas Lueder, kuris vėliau savo veikale „Kritische Geschichte der Statistik" kritikavo ne tik universitetinės statistikos atstovus, bet neigė ir pačią statistiką. Tiesą pasakius, vad. universitetinė statistika mokslo atžvilgiu ir išsyk stovėjo ne per daug aukštai, o Lueder'io laikais buvo pavirtus iš vienos pusės į valdovų politikos tarnaitę, iš antros — mėgino vaidinti žynio-pranašo rolę. O tai ir davė medžiagos Lueder'io kritikai.

Patys universitetinės statistikos atstovai savo tarpe karštai ginčijosi pirmiausia dėl klausimo, koki valstybės „ypatingumai" turi būti vadinami statistika, o koki ne. Tai daryta priklausomu nuo paties žodžio — „statistikos" — sąvokos ir jos kilmės. Dėl jos taip pat nebuvo vienos nuomonės. Visi tie, kurie tarė, kad statistika pareina nuo stato — valstybė, norėjo ir pačiai statistikai duoti daugiau politikos mokslo žymių. Kiti gi, kurie sakė, kad statistika pareina iš status — padėtis, teigė, kad statistika turi nušviesti tik esamą padėtį, nesileisdama į politikos samprotavimus ir nedaryt jokių sprendimų dėl ateities. Be to, kaip visur taip ir čia, atsirado vadinamas „aukso vidurys", kuris mėgino abidvi sąvokas ir iš jų einamas išvadas sujungti į vieną ir tuo būdu statistiką padaryti „valstybinės padėties mokslu".

Nežiūrint viso to nuomonių skirtumo, visi universitetinės statistikos atstovai buvo vienos nuomonės dėl skaitmenų, kurių jie vengte vengė. Tai ypač paaiškėjo po to, kai tūli universitetinės statistikos atstovai, pildydami savo viešpačių pageidavimus, pradėjo naudoti archyvų skaitmeninius duomenis (ypač iš gyventojų ir ūkio statistikos) ir iš jų sudarinėti st. lenteles (tabeles). Čia ištikimieji universitetinės statistikos atstovai įžiūrėjo tiesiog statistikos profanaciją, st. lentelių autorius vadino „tabelių bernais", o jų statistiką laikė žema — „gemeine", kuri visai negali būti lyginama su „tikrąja" arba „aukštąja" statistika — be skaitmenų.

14. Graunt'o „Natūra" ir Suessmilch'o „Dieviškoji santvarka".

Bet grįškim vėl prie istorijos. Dar anksčiau negu Vokietijoj Conring, Achenwall ir kt. savo „valstybinius ypatingumus" pradėjo vadinti statistika ir dėl jos ginčytis, Anglijoj (1662 m.) pasirodė nedidelis, bet savo turiniu ir forma tiems laikams visai nepaprastas veikalas vardu: „Natural and Political Observations, Mentioned in a folloving Index, ant made upon the Bills of Mortality. By Capt. John Graunt, Citicen of London — with reference to the Government, Religion, Trade, Growth, Air, Diceases, and the several Chanches of the said City".

Vėliau buvo kilęs ginčas dėl šito veikalo tikrosios autorystės, kuri buvo norima suteikti Graunt'o draugui ir vad. politinės aritmetikos vardo išradėjui W. Petty. Dėl šito dalyko net toks rimtas ir bešališkas autorius kaip C. H. Hull galutinai išsitarti nedrįsta: nors ir linksta daugiau prie nuomonės, kad tą veikalą parašęs Graunt, ne Petty, vis dėlto įdėjo tą rašinį į savo redaguotą W. Petty ekonominių rašinių rinkinį, — tiesa, su pastaba, kad to rašinio autorius yra greičiausia ne Petty, o Graunt[9].

Mums čia svarbi ne tiek pati autorystė, kiek rašinio turinys ir forma. Jau antraštė rodo, kad autorius ima Londono

---

[9] *C. H. Hull,* The Economic Writings of Sir William Petty together with the Observations upon the Bills of Mortality more probable by Captain John Graunt. Cambridge. 1899.

miesto mirčių statistikos duomenis ir rodo juos įvairiausiais atžvilgiais — politikos, tikybos, pramonės, ligų, amžiaus ir net klimato arba, kaip jis sako, oro atžvilgiu. Pažiūrėjus į patį rašinį iš tikrųjų randi visą eilę lentelių bei aiškinimų ne tik apie mirties atsitikimus paviršutinėmis jų žymėmis arba tariamomis priežastimis, bet ir priežastingumo arba pareinamybės santykiavimų ieškojimo. Mirties atsitikimuose autorius mano radęs tą politinį ir „natūrinį" dėsningumą, kurį kiek vėliau (1765 m.) jo pasekėjas Suessmilch Vokietijoj pavadino „Goettliche Ordnung" — dieviškąja santvarka. „Štai, — sako Suessmilch — išmintingiausias pasaulio Sutvertojas ir jojo Valdovas prieš savo akis leidžia gausingajai žmonių giminės armijai iš nieko prasidėti, ir tiek, kiek Jis paskyrė, gyventi. Amžinasis leidžia mums eiti pro šalį Jojo akių kiekvienam ligi tol, kol sau paskirto tikslo pasieksim; o paskui ir vėl iš kelio išeinam. Išėjimas, pro Viešpaties akių ėjimas ir patsai nuėjimas atliekamas nuostabiai tvarkingai. Mūsų pasirodymas gyvųjų pasauly vyksta iš lėto, nesigrūdant ir tam tikrais skaičiais, kurie tiek gyvųjų armijoj, tiek su pasitraukiančiais visuomet stovi tam tikram dėsningam santyky. Trumpai prieš įžengiant į gyvųjų pasaulį kai kurie dar kartą atrenkami: tai negyvai gimusieji; bet ir tai vyksta tam tikram santykyj". Ir t. t. Bet iš kur visa tai buvo patirta?— klausia Suessmilch'as ir pats atsako: „Bažnyčių rejestrai dabar yra didžioji pagalbos priemonė tiems dėsniams surasti. Tie rejestrai jau laikomi keletą šimtų metų. Bet kas ligi Graunt'o juos panaudojo tiems dėsniams surasti? Šitas išradimas buvo taip pat galimas kaip ir Amerikos atradimas: reikėjo tiktai Kolumbo, kuris savo samprotavimuose senai žinomomis tiesomis ir žiniomis nueitų toliau negu kiti"[10].

Ir iš tikrųjų, Graunt'as paėmė paprastus parapijų archyvų skaitmeninius duomenis, juos sugrupavo ir parodė, kad žmonių mirtingumas, aplamai, taip pat mirtingumo pasireiškimai atskirose gyventojų grupėse rodo mažiau ar daugiau pastovius santykio dydžius, t. y. vyksta dėsningai. O žinant mirtingumo dėsnius esą galėtų ir valstybė ar bažnyčia šį tą padaryti. Patsai autorius dar nedrįso aiškiai apibrėžtų patarimų duoti, o tik patarė pagalvoti, ar tos žinios nebūtų reikalingos arba naudin-

---

[10] *Suessmilch,* Göttliche Ordnung, I, 52/3 ir 56/7.

gos ir kitiems, — ne vien karaliui ir jojo ministeriams (žiūr. baigiamąjį sakinį jo „observacijose").

15. W. Petty politinė aritmetika.

Aiškias išvadas iš statistinių dėsningumų padarė Graunt'o draugas William Petty, kuris pakrikštijo ir patį rašymo bei tyrinėjimų krypstį, pavadindamas jį „politine aritmetika". Iš profesijos gydytojas, William Petty buvo geras matematikas, matininkas, muzikas, vertėsi laivų statyba ir kt. Bet plačiau pagarsėjo kaip visuomenės veikėjas ir merkantilistinės politikos atstovas, tą politiką pagrįsdamas daugiau statistikos duomenimis, ne žodiniais argumentais. Šiuo atžvilgiu įdomi yra jo dedikacija, įrašyta „Politinėj aritmetikoj", įteikiant ją karaliui. Prasideda toji dedikacija šitaip: „Viešpatie, kuomet kiekvienas rūpinasi Jūsų Didybei paaukoti ką nors, kas geriausia tiktų Jums laimingai žengiant į Sostą; aš drįstu paaukoti, ką mano tėvas senai yra rašęs, kad parodyti Anglų Karūnos svorį ir josios svarbumą" (to shew the weight and importance of the English Crown). Pačioj prakalboj autorius, nupiešęs bendrais, bet labai vaizdingais, bruožais to laiko Anglijos politinę, ekonominę ir socialinę padėtį ir visuomenėj ar raštuose skelbiamas pesimistines mintis, toliau šitaip prabyla: „Šitos bendros pastabos, o ir tai, kad žmonės valgo, geria ir juokiasi taip pat, kaip jie buvo iš seno papratę daryti, mane padrąsina pamėginti, ar aš negalėčiau, patsai pasitenkindamas, ir kitiems parodyti, kad Anglijos reikalai nėra taip blogoj padėtyj.

„Metodas, kurį aš tam tikslui naudoju, dar nėra labai įprastas; nes vieton naudoti vien lyginamuosius ir labai aukštus (superlative) žodžius ar intelekto argumentus, aš... pasiryžau savo įrodinėjimus išreikšti skaitmenim, svoriu ir mastu; naudoti vien proto argumentus ir imti dėmesin tiktai tokias priežastis, kurios turi regimo pagrindo prigimtyje; palikti kitiems svarstyti visa, kas priklauso atskirų žmonių nuomonės, pažiūrų, apetitų ar palinkimų... Autoriaus paduodamieji skaičiai, svarai ir mastai, kuriais jis savo įrodymus remia, gali būti ir ne visai tikri; bet jeigu jie yra netikri, tai ne tiek, kad visai sugriautų jais paremtus argumentus, o blo-

giausiu atveju — jie yra ganėtini, kad parodytų kelią į mokslą, kurio autorius siekia“[11].

Graunto-Petty pradėtieji politinės aritmetikos metodai patraukė ir daugiau šalininkų (ypač Anglijoj). Atsirado jų net ir universitetinės statistikos tėvynėj, Vokietijoj, kur vyriausiuoju politinės aritmetikos atstovu buvo anksčiau cituotasis Suessmilch. Nuo universitetinės statistikos politinė aritmetika skyrėsi daugiausia tuo, kad pirmoji „valstybės ypatingumus“ stengėsi aprašyti žodžiais ir sąmoningai vengė sausų skaitmeninių duomenų. Politinė aritmetika mėgino viską vaizduoti skaičiais, o žodžio argumentai tarnavo tik skaičiams paaiškinti. Savo esmėje betgi vienas ir antras šios ankstybosios statistikos krypsniai buvo ne savarankiški mokslai, o tarnavo politikai — su tuo gal skirtumu, kad universitetinės statistikos atstovai daugiau tarnavo atskiriems valdovams, o politinė aritmetika taikėsi bendrajai gadynės dvasiai ir ieškojo dėsningumo, paremto prigimties įstatymais arba Dievo duotąja visuomenės santvarka. Ir taip ėjo viskas ligi XIX šimtmečio pradžios, kai abudu krypsniai susiliejo į vieną ir sudarė pagrindą moderninės statistikos mokslui, kurio įkūrėju laikomas belgas Quetelet.

16. Quetelet ir jo „Socialinė fizika“.

Ad. Quetelet yra gimęs 1796 metais Gento mieste. Jau iš mažens besimokydamas Quetelet ėmė rodyti didelių gabumų matematikoj ir astronomijoj. Prie pirmos progos jisai įsitaiso net nuosavą observatoriją. 1836 m. Quetelet kviečiamas profesoriauti ką tik įsteigtoje Ecole militaire ir dėstyti astronomiją ir geodeziją. Penketai metų praėjus, Quetelet atsistoja į priekį centrinės statistikos komisijos, kuri turėjo padaryti pirmąjį visuotinį Belgijos gyventojų surašymą. Nuo to laiko Quetelet ima vadovauti ne tik Belgijos oficialinę statistiką, bet ir statistikos mokslą aplamai, ir net tarptautinę statistikos praktiką į naujas vėžes įstato.

Tokia greita Quetelet karjera ne astronomijoj, kurios profesorium jis pakviestas buvo, o statistikoj, turi savo pagrindą.

---

[11] Žiūr. *Hull išleistą* „The Economic Writings of Sir *William Petty*, I, 239, 244 ir 245.

Dar būdamas privatinis tyrinėtojas Quetelet išleido veikalą (1835 m.), kuris tiek turiniu, tiek metodais, tiek išvadomis sukėlė didelį susidomėjimą ir iššaukė ginčus, kurie tęsėsi ištisus dešimtmečius ir nėra užmiršti dar ir šiandien[12]. Veikalas pavadintas ilgoku vardu: „Sur l'Homme et le Développement de ses Facultés, un Essai de Physique Sociale". Autorius nagrinėja žmogų ir jo gyvenimą — žmogaus darbus ir sąlygas, kuriose tie darbai atliekami. Ir nagrinėja ne bendrais posakiais ar logikos sentencijomis, bet keliais, kuriuos pramynė Graunt, Petty, Suessmilch ir kiti analytinės statistikos atstovai. Ir ką gi jis randa? Pirmiausia ne tik dėsningumą, bet ir nepaprastą pastovumą visų tų statistinių duomenų, kurie rodo žmogaus būtį ir jo pasielgimus — ne tik paprastuosius, kasdieninio gyvenimo darbus, bet ir jo intelekto bei doros pasireiškimus. Ypač autorių stebina „pastovumas, kuris per metų metus rodo tuos pačius prasikaltimus toj pačioj tvarkoj pasikartojančius ir tame pačiame santyky vykdomas (sekančias) bausmes". „Visoj eilėj rašinių — sako autorius, — aš stengiausi šitą reiškinį išaiškinti; ir kas metai aš priėjau išvadą: yra biudžetas, kuris mokamas pasibaisėtinai reguliariai — tai kalėjimų, kartuvių ir ešafotų biudžetas; ir taupumo žiūrėti reikėtų pirmiausia čionai; ir per metų metus skaičiai mano spėjimus tiek patvirtina, kad dar būtų teisingiau, jei būčiau pasakęs: yra vienas mokestis, kurį žmogus sumoka daug reguliariau, negu gamtai arba valstybės iždui — tai nusikaltimui mokamas mokestis"[13]. Berods, esą vieni žmonės elgiasi vienaip, kiti —kitaip; atskiri žmonės per visą savo gyvenimą nepapildo nė vieno nusikaltimo, o kitų gal visas gyvenimas būna tik vienas nusikaltimas. Bet „jeigu žiūrėti visos masės, tai ir moraliniai arba dvasiniai apsireiškimai atitinka paprastuosius gamtos reiškinius; ir mes prieinam šios rūšies mūsų tyrinėjimų pagrindinį dėsnį, kad juo didesnis tyri-

[12] Naujausias Quetelet „socialinėj fizikoj" iškeltųjų problemų vertinimas yra atspausdintas Revue de l'Institut International de Statistique, 1937, livr. 2., prof. *Winklerio* rašinyje „Das Problem der Willensfreiheit in der Statistik", dėl kurio išsitaria prof. *F. Weyr* to paties leidinio Nr. 4 (1937), pusl. 329—334.

[13] *Quetelet,* Sur l'homme, I, 8/9.

nėjimų (observuojamų) vienetų (individų) skaičius, tuo silpniau matomos individualinės fizinio arba moralinio pobūdžio ypatybės ir juo daugiau reiškiasi tie bendrieji veiksniai, kuriais remiasi ir laikosi visuomenė. Tik kai kuriems ypatingai genialiems žmonėms yra lemta paveikti visuomenės santvarką, o ir tai dar reikalinga laiko, kad jų įtaka pilnai pasireikštų"[14].

Šiais žodžiais Quetelet pasakė ne tik savo „socialinės fizikos" pagrindinę tezę, bet apibrėžė ir statistikos pagrindinį dėsnį, kuris matematikoj jau anksčiau buvo pavadintas „didžiųjų skaičių įstatymu". Nes vad. „prasikaltimų biudžetas" iš tikrųjų yra tik vaizdingais žodžiais apibrėžtas statistikos dėsningumas, kuris visoj pilnumoj pasireiškia tyrinėjant mases arba masinius reiškinius tada, kai įsigali vadinamas didžiųjų skaičių dėsnis.

Tačiau Quetelet rašė ne statistikos mokslo vadovėlį, o piešė visuomenės ir atskiro žmogaus gyvenimo vaizdą socialinės fizikos šviesoje. Jei visuomenės gyvenime net doros ir intelekto srityse reiškiasi aiškiai matomas dėsningumas, jei net nusikaltimai pasikartoja per metų metus reguliariai ir toj pačioj tvarkoj, tai tas rodo, kad pačioje visuomenėje turi būti priežastys, kurios, pagal didžiųjų skaičių dėsnį, pasireiškia tam tikrame nusikalstamų darbų skaičiuje. Ir čia Quetelet padarė tolimesnę jau logikos išvadą: jeigu yra visuomenėje nusikaltimų, tai — „c'est la société qui prépare le crime et que le coupable n'est que l'instrument qui l'exécute. Il en résulte que le malheureux qui porte sa tête sur l'échafaud ou qui doit finir son existence dans les prisons, est en quelque sorte une victime expiatoire de la société. Son crime est le fruit des circonstances dans lesquelles il s'est trouvé; la gravité de son châtiment en est peut-être un nouveau résultat"[15].

Tyčia išrašiau šitą Quetelet išvadą pažodžiui, nes dėl jos ypač buvo daug ginčų. Kai kuriuose statistikos mokslo raštuose apie šitą vietą kalbama taip, tarytum pats Quetelet visuomenės atsakomybę už prasikaltimus dar nebuvo gana aiš-

[14] *Ten pat,* pusl. 12.
[15] *Quetelet,* Sur l'homme, II, 325/6.

kiai pabrėžęs ir esą tik „jo mokiniai“ šitą deterministinę mintį yra išplėtę. Bet cituotieji žodžiai aiškiai rodo, kad ir pats Quetelet gana griežtai buvo pabrėžęs, jog prasikaltimą prirengia visuomenė, o prasikaltėlis yra tik nelaimingas įrankis arba priemonė, kuri paskui, kaip visuomenės santvarkos ir aplinkybių auka, turi baigti savo gyvenimą kalėjime arba ant kartuvių; pati gi bausmė esanti tik nauja tos visuomeninės santvarkos ir visų aplinkybių pasekmė.

Bet Quetelet ir čia nesustojo. Teisybė, visuomenėje kaip ir visoje gamtoje, esą visokių žmonių; taigi pačioje visuomenėje yra diegai arba priežastys ir gera ir bloga — lygiai kaip ir kiekvienoje masėje arba kiekviename masiniame reiškinyje. Bet koksai gi idealas? Ir čia Quetelet jau pasirodo statistikas. Jis mato, kad visų masinių reiškinių idealiausias yra tas, kur, einant didžiųjų skaičių dėsniu, antraeiliai, necharakteringi pažymiai eliminuojasi, nyksta, o pilnumoj pasireiškia tik pagrindinės, charakteringosios ypatybės (arba patys pagrindiniai veiksniai). Aplamai, šitas savo charakteringumu idealiausias tipas yra „vidurinis“ arba „normalinis“ dydis, t. y . vidurkis. Dėl to ir žmonių visuomenės idealiausias tipas turįs būti v i d u t i n i s ž m o g u s — homme moyen, kuriame atsispindėtų visos tautos arba visuomenės charakteringosios ypatybės. Toksai vidutinis tautos žmogus esą būtų lyg kūno gravitacijos centras ir, aplamai, atitiktų lygsvarumos bei judėjimo apsireiškimus.

Bet ar toksai vidutinio žmogaus idealas yra galimas? Autorius mano, kad galimas, ir kad civilizacija bei apšvietimas su laiku padarys tai, kad charakteringosios (vidutinės) visuomenės ir atskirų žmonių ypatybės galų gale paims viršų ir tuo būdu susidarys idealios vidutinės sąlygos ir vidutiniam žmogui gražiai ir laimingai gyventi[16].

---

[16] ...un des principaux faits de la civilisation est de resserrer, de plus en plus, les limites dans lesquelles oscillent les différents éléments relatifs à l'homme. Plus les lumières se répandant, plus les écarts de la moyenne vont en diminuant; plus, par conséquent, nous tendons à nous rapprocher de ce qui est beau et de se qui est bien“ *(Quetelet,* Sur l'homme, II, 326).

17. Pagrindinės Quetelet mokslo mintys.

Paliekant šaly Quetelet rašinių filosofinę pusę (jis yra parašęs gana daug įvairaus turinio raštų, kurių pagrindinės mintys, tačiau, yra cituotame veikale „apie žmogų"), mums čia tenka konstatuoti ir trumpai formuluoti tik tas pagrindines mintis, kurios jų autorių daro tikruoju statistikos tėvu. Pirmoj eilėj čia stovi vadinamas d i d ž i ų j ų s k a i č i ų d ė s n i s. Savo esmėje tas dėsnis statistikoj buvo pažįstamas ir anksčiau, ypač iš Suessmilcho „Dieviškosios santvarkos", kur pasakyta: „jei imti tik atskirus namus skaičiuoti, tai viename rasime gimus tiktai mergaites, kitame vien berniukus arba... labai jau nelygią proporciją", — kai iš tikrųjų 20-čiai mergaičių gema 21 berniukas. Čia aiškiai matyti, kad turėta galvoje vad. didžiųjų skaičių dėsnio veikimas. Taip pat didžiųjų skaičių dėsnio veikimas buvo pastebėtas visos eilės matematikų ryšyj su mėginimais apibrėžti tikimybes ir surasti jų teorinį pagrindimą. Šiuo atžvilgiu žinomoji J. B e r n o u l l i tikimybių teorema jau visai aiškiai pabrėžia tikimybės laipsnio pareinamybę nuo didžiųjų skaičių dėsnio veikimo. Tačiau Quetelet didžiųjų skaičių dėsnį trumpai, griežtai ir drauge kiekvienam suprantamai formulavo ir jį pagrindė statistikos duomenimis iš socialinio žmonių gyvenimo.

Dėl šios priežasties ir Quetelet vad. s t a t i s t i n i a i d ė s n i n g u m a i įgauna visai kitą prasmę ir turinį. Dar politinės aritmetikos atstovams statistiniai dėsningumai buvo savo rūšies misterija, pareinanti nuo Dievo valios arba nuo gamtos paslaptingumų, kuriuos tik drąsesnieji gamtos dėsningumais arba įstatymais vadindavo. Quetelet gi statistiniai dėsningumai terodo tik pagrindinių priežasčių arba veiksnių pakankamai plačiu mastu (didžiųjų skaičių dėsniui veikiant) pasireiškimą, o nuo tų dėsningumų nukrypimai arba vadinamos „išimtys" tėra tik antraeilių arba atsitiktinių veiksnių (arba priežasčių) regimi liudytojai. Ir vis tiek, ar Quetelet formulavo „prasikaltimų biudžeto" dėsnius, ar teigė, kad „Belgijos gyventojai savo mokesnį vedyboms sumokėjo (1825—1845) tvarkingiau (reguliariau), kaip mokestį mirčiai", — jis visuomet savo formulas ir teigimus pagrindė statistikos duomenimis bei jų analiziu, kur rodikliu buvo tikimybių teorija.

Quetelet statistinių vidurkių, kaip masės arba masinių reiškinių idealo arba ideališkai charakteringo tipo, teorija taip pat savo esmėje yra teisinga, jei ją imti paties Quetelet interpretacijoj, kuri teprileido tik tokių dydžių vidurkius, kur didžiųjų skaičių dėsnio veikimas pilnumoj pasireiškia. O Quetelet vidutinis žmogus, kaip žmonių visuomenės idealas, nors ir ginčijamas ir vargu ar pasiekiamas, bet ši vidutinio žmogaus teorija buvo vienas iš galingiausių akstinų statistinių vidurkių teorijai išplėsti ir net įvairioms filosofijos problemoms aiškinti.

Statistikos praktikai Quetelet vadovaujama įtaka yra niekeno neginčijama. Pasiėmęs vadovauti Belgijos oficialinę statistiką, jis tiek daug nauja ir praktiška įnešė statistikos darbo organizacijon, kad Belgijos pavyzdį betarpiškai, arba per Tarptautinius Statistikos Kongresus, pasekė ir kitų kraštų oficialinės statistikos įstaigos.

Aplamai, Quetelet įtaka tiek statistikos mokslui — jo teorijai ir metodologijai, tiek statistikos praktikai — statistikos darbo organizacijai ir technikai, yra labai didelė. Dėl to bus teisingiausiai, jei Quetelet moderniškojoj statistikoj paskirsime tokią pat vietą, kaip Graunt-Petty vad. politinėj aritmetikoj, t. y. jį pavadinsime moderniškosios statistikos kurėju.

18. Modernosios statistikos krypsniai.

Šių dienų statistikos mokslo atstovus ir literatūrą tenka skirstyti į dvi pagrindines grupes, žinoma, su neišvengiamu „aukso viduriu". Kaip ir kitose mokslo šakose, šios grupės yra ne tiek esminio, kiek metodologinio pobūdžio, kur induktyviniam ar deduktyviniam metodui mažesnis ar didesnis palinkimas nulemia prie vienos arba antros grupės priklausomybę. Gal tik tas skirtumas, kad statistikoj vieno ar antro metodo pasirinkimas lengviau nulemia ir turinio bei formos skirtumus. Mat, kitose mokslo šakose induktyvinio ir deduktyvinio metodo šalininkai paprastai vartoja tas pačias, arba labai artimas, išreiškimo priemones (žodinę argumentaciją) ir daugiau skiriasi tik išrodinėjimo tvarka (iš faktų arba konkretiškų atsitikimų išvedamos tezės, arba išstatytoms, aprioristinėms tezėms patvirtinti tiekiami faktai bei konkretiški duo-

menys). Statistikoj, kur operuojama skaičiais arba skaitmeniniais duomenimis, deduktyvinio metodo šalininkams išreiškimo priemonės peršte peršasi iš matematikos su jos gausingais ir parankiais skaičių bei reiškinių simboliais. Dėl šios priežasties modernioji statistika net vadinama dažnai matematine statistika — ir jos priklauso daugiausia deduktyvinio metodo statistikoj šalininkai, ir istoriškai aprašomąja, kurios priklauso induktyvinio metodo šalininkai.

Dėl istorinių aplinkybių, matematinės ir istoriškai aprašomosios statistikos krypsnių atstovų daugumą galima skirstyti nagrinėjimų klausimų rūšimi ir net geografiniais pažymiais. Prancūzijos ir Anglijos, taip pat Amerikos J. V. ir iš dalies Italijos bei Rusijos, statistikų dauguma, su mažomis išimtimis, daugiau buvo veikiama Quetelet tikimybiškai matematinės argumentacijos. Dėl to ir jų statistikos mokslo literatūra ne tik gausinga matematikos teikiamomis statistinio analizio bei vaizdavimo formomis, bet ir apsirėžia daugiausia statistinio analizio problemomis. Iš kitos pusės nemaža statistikų, ypač Vokietijoj ir vokiečių kultūros įtakoj esančiuose kraštuose, buvo veikiama Muencheno universiteto profesoriaus ir Bavarijos oficialinės statistikos vadovo Georg von Mayr, kuris pavyzdingai sutvarkė Bavarijos oficialinę statistiką ir parašė didžiulį trijų tomų veikalą vardu „Statistik und Gesellschaftslehre“ (neskaitant daugybės smulkesnių raštų) istoriškai aprašomuoju metodu. G. v. Mayr buvo griežtas vad. matematinės statistikos priešininkas, ir tokiais jis išauklėjo mažne visą Vokietijos statistikų kartą, — nežiūrint į tai, kad moderniškosios matematinės statistikos atstovai daug daugiau remiasi Vokietijoj „nepripažintų“ Bortkiewicz'o ir Lexis'o statistikškai matematiniais tyrinėjimais, negu paties Quetelet raštais. Pastaraisiais laikais ir Vokietijoj vad. matematinės statistikos šalininkų jau daugiau atsiranda.

Šalia šių dviejų griežtai priešingų ir kraštutinių krypsnių visuomet buvo vad. aukso vidurys, kur principialiai neneigiamas nė vienas metodas ir su nauda pasikliaujama abiejų krypsnių tiekiamomis statistinio stebėjimo, nagrinėjimo (analizio) bei vaizdavimo formomis. Ir galima net pastebėti, kad pastaraisiais laikais vis daugiau ir daugiau pastaroji statistikų srovė įsigali.

D. Statistikos darbo organizacija.

19. Statistikos darbo organizacijos principai.

Visas statistikos darbas galima suskirstyti į tris etapus arba stadijas: a) statistinis stebėjimas, t. y. statistikos duomenų rinkimas, b) statistinis vaizdavimas, t. y. surinktųjų arba turimųjų statistikos duomenų klasifikacija (statistinės medžiagos apdirbimas), ir c) statistinis analizis, t. y. statistikos duomenų nagrinėjimas. Visi šie trys statistinio darbo etapai techniškai organizaciniu atžvilgiu gali būti išdalinti, bet gali būti ir vieno statistikos organo (įstaigos) arba net vieno asmens vykdomi. Šiais laikais, tačiau, kai statistika apima mažne visas žmonių socialinio ir ekonominio gyvenimo sritis, vadinamajai privatinei statistikai, t. y. privatiniams statistikos tyrinėtojams, statistinio stebėjimo srityje vis mažiau ir mažiau progos pasireikšti telieka. Patsai vienas moderniosios statistikos metodais statistinis stebėjimas yra tiek platus ir tokios daugybės darbo reikalingas, kad jį atlikti tegali tik lėšomis ir žmonėmis gausiai aprūpinti specialiniai statistikos organai arba įstaigos.

Juridiniu atžvilgiu, statistikos darbo organizacija gali būti oficialinė ir privatinė. Sulig tuo kalbama ir apie oficialinę ir privatinę statistiką. Patys žodžiai rodo, kad pirmosios rūšies statistikos darbas turi būti vykdomas oficialinių, t. y. viešosios teisės (valstybės, savivaldybinių ar kt.) organų, o antrosios — privatinių įstaigų, organizacijų ar šiaip asmenų. Yra dar ir kitas skirtumas. Kai statistikos medžiaga yra renkama privatinių statistikos įstaigų (ar asmenų), tai piliečiai gali prašomų statistikos duomenų suteikti arba nesuteikti. Tačiau, kai statistikos medžiaga renkama oficialinių statistikos organų, tai pilietis, paprastai, jau įstatymo keliu yra įpareigotas statistikai reikalingus duomenis suteikti, o priešingu atveju — jis gali būti baudžiamas. Ir dabartiniu metu Lietuvoj veikiantis „Valstybės statistikos įstatymas“ § 14 nustato: „asmuo, atsisakęs be teisėtos priežasties suteikti valstybės statistikos reikalui žinių, arba suteikęs žinomai melagingų žinių, yra baudžiamas bauda ligi 250 litų“.

Sistemos atžvilgiu, statistikos darbas gali būti organizuotas centralizacijos arba decentralizacijos principu. Šie du principai šiandien kai kuriose valstybėse, arba net toj pačioj valstybėj atskiruose statistikos sektoriuose, dar konkuruoja. Tačiau vis daugiau ir daugiau įsigali centralizacijos principas statistinio darbo organizacijose. Tai ir suprantama. Svarbiausioji statistikos, kaip masių arba masinių reiškinių vaizdavimo bei matavimo priemonė yra statistikos duomenų klasifikacija charakteringomis žymėmis ir pažymiais tuo būdu, kad šitie duomenys galėtų būti lyginami esminėmis, erdvinėmis ar laiko ypatybėmis. Tačiau turimųjų statistikos duomenų lyginimas yra galimas tik tuo atveju, jei jie yra gauti tais pačiais statistinio stebėjimo metodais ir, aplamai, gali būti laikomi statistinio lyginimo atžvilgiu vienalyčiais. Jei gi statistinis stebėjimas arba medžiagos apdirbimas vykdomas autonominių organų, tai apie statistinį vienalytiškumą kalbos būti negali. Dėl to statistikos darbo organizacijoj vis daugiau įsigali vad. centralizacijos principas. Tai nereiškia — pagaliau, ir nereikalinga, — kad savarankiškų statistikos organų visai būti negali arba kad visos socialekonominio gyvenimo sritys statistikos stebimos ir vaizduojamos turėtų būti tiktai vieno statistikos organo. Tokia griežta, be jokių išimčių vykdoma, statistikos darbo centralizacija būtų ir pačiai statistikai nenaudinga. Nes yra tokios rūšies masių arba masinių reiškinių, kur statistiškai prasmingas darbas gali būti vykdomas tik pilna to žodžio prasme tos rūšies specialistų. Jei šiandien kalbama apie statistikos darbo organizacijos centralizaciją, tai tik ta prasme, kad bet kurios rūšies statistiniai stebėjimai vykdomi su centrinio statistikos organo žinia, ir jo organizuojami taip, kad atitiktų statistikos teorijos bei metodologijos reikalavimus; be to, mažiau ar daugiau savarankiški arba specialiniai statistikos organai įpareigojami savo turimą statistikos medžiagą klasifikuoti bendraisiais statistinės klasifikacijos dėsniais ir ją patiekti centriniam statistikos organui. Tose gi srityse, kur paraleliškųjų statistikos organų veikimas bei darbas neišvengiamas (atskirų valstybių statistikos organai, toj pačioj valstybėj atskirų autonominių vienetų arba, aplamai, viešosios teisės organų paraleliniai statistikos organai) — bent stengiamasi, kad atskiruose socialinio ar ekonominio gyvenimo sektoriuose

statistinis stebėjimas vyktų kiek galint vienodesniais principais, pagal vienodą bendraisiais bruožais nomenklatūrą ir kad statistikos medžiagos pirminė klasifikacija būtų patiekta vienodais principais sudarytose lentelėse (tabelėse). Šios rūšies statistinio darbo koordinacija atskirose valstybėse paprastai vyksta centriniam arba oficialiam statistikos organui vadovaujant, o tarptautinėje plotmėje tarptautinių statistikos organų (tarptautinio statistikos instituto ir kt.) arba specialinių statistikos suvažiavimų (kongresų) nutarimų arba net ir tarptautinių statistikos konvencijų nuostatais.

20. Statistikos darbo organizacija Lietuvoje.

a) Statistikos darbo organizacija prieš D. karą.

Nieko tikresnio tuo tarpu negalima pasakyti apie statistikos darbo organizaciją D. Kunigaikščių Lietuvoje, nors statistinio pobūdžio surašinėjimai, kaip istorinėj dalyj buvo parodyta, tais laikais buvo daromi. Sprendžiant iš anologijos reikėtų manyti, kad specialinių statistikos organų šiandienine prasme D. Kunigaikščių Lietuvoje nėra buvę. Šios rūšies statistikos organai iš viso teatsiranda tik naujaisiais laikais su moderniškosios statistikos metodų įsigalėjimu.

Prieškarinė Lietuva savo specialinių statistikos organų nei specialinių statistikos surašinėjimų nėra turėjusi, nežiūrint į tai, kad sistematingų statistikos surašinėjimų Lietuvoje buvo daroma nuo 1878 m. Tais metais, būtent, buvo padarytas per administracijos organus savo rūšies žemės ūkio surašymas Rusijoj ir Lietuvos gubernijose. Sekantis panašus surašymas yra buvęs 1888 m. ir statistiškai metodologiniu atžvilgiu tobulesnis — 1905 m. Be to, 1908 m. yra buvęs savo rūšies pramonės statistinis surašymas. Jau specialinių statistikos organų gana gerai organizuotas buvo visuotinis gyventojų surašymas 1897 m. Savo keliu kiekvienoje gubernijoje prie gubernatoriaus kanceliarijos veikė vad. gubernijų statistikos komitetai,kurie per administracijos organus rinkdavo įvairias žinias iš gyventojų ir ūkio sričių, skelbdavo jas vad. gubernijų „Primenamose knygelėse“ („Pamiatnaja knižka ... gub.“) ir patiekdavo duomenis gubernatoriui metiniams raportams, kurie būdavo patiekiami carui.

Išskiriant 1897 m. visuotinį gyventojų surašymą, kuris visoje Lietuvoje (be Klaipėdos krašto) buvo vykdomas pagal centrinio Rusijos statistikos komiteto (prie vidaus reikalų ministerijos Petrapilyj) programą ir jos direktyvas, organizaciniu atžvilgiu prieškarinės Lietuvos socialekonominių reiškinių statistinis stebėjimas ir vaizdavimas vyko: a) Kauno, Vilniaus gubernijose ir kt. rajonuose pagal centrinio Rusijos komiteto direktyvas Petrapilyje, b) Suvalkų gubernijoj pagal centrinės statistikos komisijos direktyvas Varšuvoj ir c) pagal Prūsijos Statistikos valdybos direktyvas (Berlyne) Klaipėdos krašte. Dėl šios priežasties beveik nesiduoda lyginami ir statistikos duomenys prieškarinėj Lietuvoj.

b) Statistikos darbas po 1918 m.

Susikūrus Nepriklausomai Lietuvos valstybei po Didžiojo karo 1919 m. rugsėjo 6 d. tuo laiku buvusio prekybos ir pramonės ministerio prof. Šimkaus įsakymu prie sakytos ministerijos buvo sudarytas bendrosios statistikos departamentas, kuriam buvo numatyti šie skyriai: 1) bendroji statistika, 2) finansų statistika, 3) prekybos ir pramonės st., 4) gyventojų st. ir, be to, 5) raštinė. Tačiau, šiam statistikos departamentui platesniu mastu statistikos darbą suorganizuoti Lietuvoje nepavyko. Iš jo veikimo darbų yra užsilikęs tiktai kompiliatyvinis statistikos duomenų leidinys vardu „Statistinės žinios apie Lietuvą“. Bet kurios rūšies savarankiškų statistikos stebėjimų statistikos departamentas nėra padaręs. Savo keliu, maždaug tuo pačiu laiku besiorganizuojant kitoms ministerijoms, buvo sudaryti specialiniai statistikos skyriai Susisiekimo ministerijoj ir Žemės Ūkio ministerijos miškų ir žemės ūkio departamentuose. Susisiekimo ministerijos geležinkelių valdybos ir Miškų departamento statistikos skyriai sistematizavo ir klasifikavo sakytose įstaigose esamus ir nuolat gaunamus skaitmeninius duomenis tų pačių įstaigų vadovybėms orientuotis. Tuo tarpu prie Žemės Ūkio departamento esantis statistikos skyrius turėjo ypatingai daryti sistematiškus stebėjimus apie pasėlių stovį, derlius, taip pat žemės ūkio produktų kainas. Be to, maždaug

tuo pačiu laiku muitinėse kaupėsi statistikos duomenų apie užsienių prekybą, kurie prašyte prašėsi apdirbami ir skelbiami. Taip pat prie Žemės Ūkio Ministerijos veikianti Žemės Reformos Komisija darė savarankiškus statistikos stebėjimus ir surinktos medžiagos klasifikavimus apie dvarų žemes, taip pat apie bežemius ir mažažemius.

Žodžiu, pirmaisiais Lietuvos Nepriklausomybės metais statistikos darbas Lietuvoje buvo jei ne deorganizuotas, tai bent decentralizuotas. Šitam darbui sutvarkyti ir reorganizuoti centralizacijos principu 1920 m. gegužės m. 17 d. jau Finansų, Prekybos ir Pramonės ministerio įsakymu buvo sudaryta tarpžinybinė visų ministerijų atstovų komisija statistikos darbui reorganizuoti. Ši komisija išdirbo valstybinės statistikos įstaigos organizacijos projektą, pagal kurį buvo įsteigtas centralinis statistikos biuras ir nuolatinė statistikos komisija. Kadangi nuo šio momento Lietuvos statistinio darbo organizacija priėmė formas, užsilikusias ligi šiol, tai čia patiekiama tų dviejų įstaigų statutiniai tekstai, kurie buvo paskelbti „Vyriausybės Žiniose" 1921 m. gegužės m. 21 d. „Vyriausybės Žinių" Nr. 65, eil. Nr. 587 ir 588.

Centralinis Statistikos Biuras.

§ 1. Centralinis statistikos biuras steigiamas statistikos darbams vykdyti.

§ 2. Centralinis statistikos biuras yra vyriausias Valstybinės statistikos organas ir dirba Finansų, Prekybos ir Pramonės Ministerijos prižiūrimas.

§ 3. Centralinio statistikos biuro tikslas yra:

a) parengti planus statistikos srities darbams atlikti ir įnešti į nuolatinę statistikos komisiją peržiūrėti ir galutinai suredaguoti;

b) sekti statistikos darbą ir mokslą užsienio literatūroje ir pritaikinti Lietuvos Valstybinei statistikai tikslingiausius jos metodus;

c) rinkti, patikrinti ir tvarkyti statistikos medžiagą, skelbti atskiras žinias spaudoje ir leisti periodinius statistikos rinkinius;

d) informuoti statistikos daviniais Valstybines įstaigas;

e) ministeriui reikalaujant, atlikti atskirus darbus.

§ 4. Centralinio statistikos biuro parengtais ir nustatytu keliu patvirtintais planais privalo vaduotis visi statistikos skyriai.

Pastaba. Kiekviena įstaiga, norėdama susirinkti statistikos žinias, turinčias bendros reikšmės, prieš vykdymą pristato šių darbų planus centraliniam statistikos biurui peržiūrėti.

§ 5. Centralinis statistikos biuras susideda iš šių skyrių: 1. Suvestinių skyrius, 2. Leidimų skyrius, 3. Raštinės ir informacijos dalis.

§ 6. Centralinio statistikos biuro direktorius arba jo pavaduotojas pirmininkauja nuolatinėje statistikos komisijoje.

D-ras *K. Grinius*
Ministeris Pirmininkas.

J. Dobkevičius
L. e. Finansų, Prekybos ir Pramonės
Ministerio pareigas.

Kaunas, 1921 m. gegužės 17 d.

588.

Nuolatinė statistikos komisija.

§ 1. Statistikos žinioms rinkti ir tvarkyti, rinkimo ir tvarkymo būdams nustatyti ir įvairiose įstaigose statistikos darbų sistemai suvienodinti prie centralinio statistikos biuro steigiama nuolatinė statistikos komisija.

§ 2. Nuolatinei statistikos komisijai priklauso šie darbai:

a) periodinius sąrašus ir kitus statistikos darbus svarstyti,

b) peržiūrėjimas ir galūtinės redakcijos nustatymas visų planų statistikos medžiagai rinkti ir metodams jos sutvarkymui nustatyti,

c) svarstymas projektų dėl sudarymo tam tikrų statistikos aparatų įvairiose įstaigose ir statistikai reikalingų specialistų ruošimas.

§ 3. Nuolatinės statistikos komisijos nutarimus jos pirmininkas duoda patvirtinti tam ministeriui, kurio darbo sritis tuo nutarimu paliesta. Parašytosios tuo reikalu rezoliucijos pranešama komisijos žiniai.

§ 4. Nuolatinės statistikos komisijos sudėtis yra ši: Pirmininkas — Centralinio statistikos biuro direktorius, Vice-pirmininkas — Centralinio statistikos biuro direktoriaus pavaduotojas, Sekretorius — Centralinio statistikos biuro korespondencijos vedėjas. Atskirų ministerijų statistikos skyrių vedėjai, arba tam tikslui paskirti į komisiją žmonės ir pakviesti žmonės-specialistai.

§ 5. Nuolatinė statistikos komisija susirenka centralinio statistikos biuro rūmuose, pirmininkui juos pakvietus.

§ 6. Dalyvavimas komisijos posėdžiuose laikomas tarnybos pareigų ėjimu ir privalomas visiems komisijos nariams.

§ 7. Komisijos posėdžiai yra teisėti, jei juose dalyvauja, neskaitant pirmininko, ne mažiau kaip pusė visų narių. Nutarimai daromi balsų dauguma. Balsams pusiau pasidalinus nusveria pirmininko balsas.

§ 8. Komisijos nariai atskiro atlyginimo už darbą negauna.

D-ras *K. Grinius*
Ministeris Pirmininkas.

J. Dobkevičius
L. e. Finansų, Prekybos ir Pramonės
Ministerio pareigas.

Kaunas, 1921 m. gegužės 17 d.

Iš patiekiamųjų nuostatų matyti, kad centralinis statistikos biuras buvo numatytas technikiniu statistikos darbo organizacijoj organu, kuris privalo veikti vadinamos nuolatinės st. komisijos derektyvomis arba atskiram atvejui apibrėžiamais nuostatais. Komisija, kiek matyti iš § 4, turėjo susidėti iš trijų centralinio statistikos biuro aukštesniųjų valdininkų ir „atskirų ministerijų statistikos skyrių vedėjų arba tam tikrų paskirtų į komisiją žmonių ir pakviestųjų specialistų". Vadinasi, iš karto buvo numatyta, kad atskirose ministerijose ar įstaigose esantieji st. skyriai nebus likviduoti, tik privalės derintis prie šios komisijos nuostatų tiek, kiek tai liečia statistikos darbų sistemą. Bet kuri įstaiga, tačiau, norėdama rinkti kurios nors rūšies st. žinias, privalėjo šiam reikalui paruoštą projektą ir darbų planą pristatyti centralinės komisijos biurui ir komisijai patvirtinti.

Taip buvo numatyta st. darbo organizacijos nuostatuose. Statistikos praktikoj centralinis statistikos biuras, vos spėjęs dorai susiorganizuoti, pirmiausia pradėjo rūpintis Lietuvos gyventojų visuotiniojo surašymo paruošimu, taip pat užsienių prekybos ir kainų statistikos organizavimu. Šie darbai, ir ypač visuotiniojo gyventojų surašymo paruošimas bei jo pravedimas, tiek užėmė laiko, kad bet kokiomis direktyvomis ar instrukcijomis kitų įstaigų st. darbams vykdyti nebuvo laiko. Juo labiau, kad nuo 1923 m. spalių m. 3 d. Finansų ministerio įsakymu centralinis st. biuras perėmė iš Žemės Ūkio Ministerijos ir visą vad. žemės ūkio statistikos skyrių su visais jo tarnautojais, korespondentais ir kt. Taip pat nuo 1924 m. pra-

džios centralinis statistikos biuras betarpiškai pradėjo rinkti žinias muitinėse apie užsienių prekybą, ir tuo būdu koncentravo užsienių prekybos statistiką. Kadangi gyventojų natūrinio judėjimo statistika jau anksčiau buvo prie st. biuro sukoncentruota, tai galima sakyti, kad nuo 1924 m. visas statistikos darbas buvo sukoncentruotas st. biure, išskiriant, tačiau, Žemės Ūkio Ministerijoj Miškų Departamente miškų st., Susisiekimo M-joj susisiekimo st. ir kriminalinę bei intelektualinę statistiką, kur duomenys vis dar be ypatingos sistemos kaupėsi atitinkamose įstaigose (Teisingumo ir Švietimo Ministerijose).

c) Statistikos darbo organizacija Klaipėdos krašte.

Klaipėdos kraštui prisijungus prie Lietuvos, teko rūpintis statistikos darbo suderinimu abiejose Lietuvos dalyse. Patsai Klaipėdos kraštas nuo 1920 m. prie Direktorijos turėjo savo statistikos biurą, kurio kompetencijoj buvo statistikos duomenų rinkimas, jų apdirbimas ir skelbimas Klaipėdos krašte. 1925 m. spalių 1 d. tarp Finansų Ministerijos ir Klaipėdos krašto Direktorijos buvo sudaryta sutartis, pagal kurią prie Direktorijos esantis st. skyrius perorganizuojamas į st. biurą Lietuvos centralinio statistikos biuro pagrindais. Klaipėdos st. biurą ir jos tarnautojus apmoka Klaipėdos krašto Direktorija, tačiau, padarytąja sutartimi, 35% visų to biuro išlaidų apmoka centro valdžia. Be to, sutartimi numatyto ir patvirtinto statuto § 15 sako: „atskiriems darbams atlikti pašalpa gali būti padidinta centro valdžiai ir Klaipėdos krašto Direktorijai tam reikalui susitarus".

Šia sutartimi tarp kita ko Klaipėdos krašto Direktorija sutiko, kad:

1) Klaipėdos krašto statistika būtų suderinta su Did. Lietuvos st.,

2) Klaipėdos krašto st. biuras privalo atlikti st. darbus Centraliniam St. Biurui pareikalavus,

3) Klaipėdos krašto st. biuro direktorius yra atsakomingas ne tik prieš Direktoriją, bet ir prieš Centr. St. Biurą, nes jisai „skiriamas ir atleidžiamas Direktorijos Lietuvos centraliniam st. biurui sutikus",

4) Metinę Klaipėdos krašto st. biuro sąmatą tvirtina Klaipėdos krašto Direktorija susitarus su Centr. St. Biuru. Be to, tuo pačiu susitarimu Klaipėdos krašto statistikos biuro direktorius įeina nariu į nuolatinę st. komisiją prie Centr. St. Biuro. Šios gi komisijos nutarimai, kai jie Finansų Ministerio patvirtinti, yra jau privalomi ir Klaipėdos krašto st. biurui.

Svarbesnieji Klaipėdos krašto st. biuro statuto 1925 m. spalių 1 d. Finansų Ministerio patvirtinti nuostatai yra šie:

Klaipėdos krašto St. Biuro statutas.

1. Klaipėdos krašto Direktorijos Statistikos Skyrius reorganizuojamas į Klaipėdos krašto Statistikos Biurą.
2. Klaipėdos krašto Statistikos Biurui priklauso visi Klaipėdos krašto statistikos darbai ir reikalai, būtent:

I. Demografija.

a) periodiniai bei ekstra gyventojų surašymai ir suskaitymai,
b) natūralus gyventojų judėjimas (gimimai, mirimai, vedybos),
c) mechaninis gyventojų judėjimas (ateivybė ir išeivybė),
d) sanitarinė statistika.

II. Ekonominė statistika.

a) pramonės įmonės ir jų produkcija,
b) kainų statistika,
c) pragyvenimo minimumo statistika,
d) darbo rinka,
e) streikų statistika.

III. Finansų statistika.

a) Klaipėdos krašto finansai,
b) bankų ir kredito įstaigų statistika,
c) protestuotų vekselių statistika,
d) konkursų statistika,
e) akcinių bendrovių statistika,
f) taupomųjų kasų statistika.

IV. Žemės ūkio statistika.

a) žemės plotų naudojimas,
b) pasėlių stovis bei derlius,
c) naminių gyvulių suskaitymas,
d) užkrečiamųjų naminių gyvulių ligų statistika,
e) žemės ūkio bendrovių statistika.

V. Administracijos statistika.

a) švietimo darbas,
b) gaisrų statistika,
c) kriminalinė statistika,
d) municipalinė statistika,
e) rinkimų statistika.

3. Klaipėdos krašto Statistikos Biuras atlieka ir kitus, be paminėtų 2 p., darbus, taip savo iniciatyva, taip pat Klaipėdos Krašto Direktorijai bei Lietuvos Centraliniam Statistikos Biurui pareikalavus.

4. Klaipėdos krašto statistikos darbas suderinamas su Lietuvos statistikos darbu. Tam reikalui kiekvienam atskiram darbui Klaipėdos krašto Statistikos Biuras sustato planą, kuris visų pirma patiekiamas Direktorijai apsvarstyti. Planas su Direktorijos nuomone pristatomas Lietuvos Nuolatinės Statistikos Komisijos svarstymui.

Jei Nuolatinė Statistikos Komisija planui pritaria, tai jis pristatomas galutiniam p. Finansų Ministerio patvirtinimui. Taip pat p. Ministeriui priklauso išsprendimas nuomonių skirtumo.

5. Kiekvienas statistikos planas turi susidėti iš:

a) tam tikroms žinioms rinkti formuliarų su formuliaro kiekvieno punkto paaiškinimu,
b) žinioms rinkti instrukcijų bei taisyklių,
c) apdirbimo tvarkos su lentelių formomis,
d) skelbimų formų.

6. Nuolatinėj Statistikos Komisijoj savo nuomonę palaikyti Klaipėdos krašto Direktorija turi teisės paskirti, be Klaipėdos krašto Statistikos Biuro Direktoriaus, ir savo atskirą atstovą.

7. Patvirtintosios instrukcijos bei taisyklės (5 ir 6 §§) skelbiamos visuomenės žiniai „Klaipėdos Krašto Valdžios Žiniose".

8. Klaipėdos krašto Statistikos Biuras susideda iš: Biuro Direktoriaus, vieno statistiko (Direktoriaus pavaduotojo), dviejų statistikos padėjėjų, dviejų skaičiuotojų (raštininkų), vieno ekspeditoriaus.

Pastaba: be etatinių tarnautojų, atskiriems darbams gali būti kviečiami neetatiniai tarnautojai.

9. Klaipėdos krašto Statistikos Biuro Direktorius yra atsakomingas prieš Direktoriją ir Centralinį Statistikos Biurą už visus statistikos darbus asmuo. Jis yra skiriamas ir atleidžiamas Direktorijos, Lietuvos Centraliniam Statistikos Biurui sutikus.

10. Biuro Direktorius sulig savo pareiga yra ir Klaipėdos krašto Direktorijos decernentas statistikos reikalams. Biuro Direktorius įeina į Lietuvos Nuolatinę Statistikos Komisiją kaipo komisijos narys.

11. Klaipėdos krašto Statistikos Biuro Direktoriui, be gryno statistikos darbo, priklauso dar šie reikalai:
    a) paskyrimas ir atleidimas neetatinių tarnautojų, Direktorijai sutikus,
    b) paskyrimas ir atleidimas etatinių tarnautojų, einant bendra nustatyta Klaipėdos krašto tvarka, remiantis Statistikos Biuro Direktoriaus pasiūlymu,
    c) lėšų aikvojimas, einant veikiančiomis Klaipėdos krašto taisyklėmis,
    d) Klaipėdos krašto statistikos atstovavimas visose įstaigose,
    e) visi kiti priklausantieji savistoviam viršininkui reikalai.

    Pastaba: apie visas permainas Biuro tarnautojų sąstate turi būti pranešama Lietuvos Centraliniam Statistikos Biurui.

12. Klaipėdos krašto Statistikos Biuro Direktorius gauna atlyginimą einant X—XI Klaipėdos krašto tarnautojų kategorija.
13. Kiti etatiniai tarnautojai gauna atlyginimą einant sekančiomis Klaipėdos krašto kategorijomis:

Vyr. Sekretorius	— kategorija	7—8
2 Sekretorius	„	6—7
2 Asistentas	„	5—6
1 Padėjėjas	„	3—4

14. Klaipėdos krašto Statistikos Biuro tarnautojai naudojasi visomis privilegijomis bei pašalpomis, teikiamomis kitiems krašto tarnautojams.
15. Klaipėdos krašto Statistikos Biuras gauna lėšas iš Klaipėdos krašto Direktorijos, einant patvirtinta sąmata. Klaipėdos krašto Direktorija pasiima visas išlaidas krašto Statistikos Biurui išlaikyti. Klaipėdos krašto Direktorija gauna iš Centro Valdžios pašalpą krašto Statistikos Biurui išlaikyti 35% bendro jos biudžeto. Atskiriems darbams atlikti pašalpa gali būti padidinta Centro Valdžiai ir Klaipėdos krašto Direktorijai tam reikalui susitarus.

    Pastaba I. Sulig biudžeto Klaipėdos krašto Direktorija turi suteikti krašto Statistikos Biurui butą, kūrą, šviesą, bei sargą.

    Pastaba II. Lėšos teikiamos krašto Statistikos Biuro užlaikymui neatskaitomos einant Klaipėdos krašto Konvencijos 35 §, iš bendrų lėšų, skiriamų Klaipėdos kraštui.

16. Metinę Klaipėdos krašto Statistikos Biuro sąmatą tvirtina Klaipėdos krašto Direktorija, susitarus su Centraliniu Statistikos Biuru.
17. Klaipėdos krašto Statistikos Biuras turi atskirą antspaudą su savo užvardymu.
18. Klaipėdos krašto Statistikos Biuro korespondencija siunčiama ir gaunama neapmokamai prisilaikant veikiančių Lietuvoje taisyklių.
19. Šis statutas įsigalioja ponui Finansų Ministeriui iš vienos pusės ir Klaipėdos krašto Direktorijai iš kitos pusės jį patvirtinus.

    Statuto permainų tvarka tokia pat kaip ir statuto tvirtinimo“.

d) Statistikos darbo reorganizacija 1930 m.

Centralinio Statistikos Biuro darbas vis plėtėsi, nes buvo reikalas vis naujas ir naujas socialinio ir ekonominio gyvenimo sritis statistikos metodais stebėti ir tyrinėti. Be demografijos, žemės ūkio, pramonės, darbo, kainų, užsienių prekybos ir kt. reiškinių stebėjimo ir vaizdavimo, buvo reikalas pradėti ir vad. ekonominės konjunktūros kitimų sistematišką stebėjimą ir vaizdavimą. Tam tikslui buvo reikalinga ne tik duomenų, kurie buvo galima gauti vienkartiniais surašymais, bet ir nuolatos vienu ar kitu metodu stebėti ūkinio gyvenimo kitimus ir tam reikalui rinkti betarpiškai statistikos duomenis iš prekybos, pramonės ir kt. verslo įmonių bei asmenų. Tačiau įstatyminio pobūdžio nuostatai, kuriais buvo paremtas Centralinio St. Biuro nuolatinis st. komisijos darbas, nenumatė nei šioms įstaigoms teisės reikalauti statistikos duomenų iš piliečių bei įstaigų, nei piliečiams prievolės tas žinias statistikos reikalui teikti. Taip pat ir statistinio darbo centralizacijos principas nepakankamai aiškiai buvo apibrėžtas, ir kai kurios įstaigos vėl mėgindavo savarankiškus statistikos surašinėjimus daryti, dažnai neatsižvelgdamos nei statistinės metodologijos reikalavimų, nei būtino reikalo. Šiais ir dar kitais sumetimais 1929 m. buvo paruoštas valstybės statistikos įstatymo projektas, kuriuo Centralinis Statistikos Biuras padaromas vyriausiu visų valstybinės statistikos darbų vykdomuoju organu, o vietoj buvusios nuolatinės komisijos sudaroma Statistikos Taryba, kurios du nariai imami iš centralinio st. biuro (direktorius ir redaktorius), vienas Klaipėdos krašto direktorijos atstovas, po vieną atstovą nuo V. D. universiteto, Žemės Ūkio Rūmų ir Prekybos ir Pramonės Rūmų ir po vieną atstovą iš tų ministerijų, kurios yra arba gali būti suinteresuotos tarybos posėdyje nagrinėjamo klausimo išsprendimu. Šiuo įstatymu taip pat numatyta, kad „visi asmenys ir įstaigos turi, statistikos biurui pareikalavus, teikti valstybės statistikos reikalui žinių“, o už neteikimą arba už melagingų žinių teikimą numatomos sankcijos (pabauda ligi 250 litų). Pačiam Centraliniam St. Biurui duota teisė pateikiamas žinias tikrinti ir tam reikalui „apžiūrėti atitinkamus dokumentus, bustus ir įrengimus“ (§ 4).

Taip pat bet kurios viešosios teisės įstaigos, kurios renka statistikos žinias, turi, Centraliniam St. Biurui pareikalavus, tų žinių jam pristatyti.

Aplamai, naujuoju įstatymu statistikos darbo organizacija Lietuvoj žymiai sustiprinta centralizacijos kryptimi. Patsai įstatymas ministerių Kabineto buvo priimtas 1930 m. birželio mėn. ir jo tekstas paskelbtas „Vyriausybės Žiniose“ Nr. 332, eilės Nr. 2272. Jo tekstas seka:

1.

Valstybės statistiką visoje valstybėje veda Centralinis Statistikos Biuras.

Kitos viešosios teisės įstaigos turi teisę vesti statistiką savo darbo srityje.

2.

Centralinis Statistikos Biuras:

1) renka, tikrina ir apdirba įstatymų arba vyriausybės reikalaujamus statistikos duomenis;
2) daro įstatymų reikalaujamus visuotinius gyventojų, žemės ūkio ir kitus surašymus;
3) skelbia valstybės statistikos duomenis ir leidžia periodinius statistikos biuletenius ir metraščius ir kitokius statistikos leidinius;
4) vienodina visos valstybės viešosios teisės įstaigų dirbamą statistikos darbą ir suveda tų įstaigų statistikos duomenis;
5) bendradarbiauja su tarptautinėmis ir kitų valstybių statistikos įstaigomis.

3.

Centralinis Statistikos Biuras dirba valstybės statistikos darbus visoje valstybėje betarpiškai. Tačiau jis gali pavesti šituos darbus atlikti savo nustatyta tvarka kitoms viešosios teisės įstaigoms, su jomis susitaręs.

4.

Centraliniam Statistikos Biurui pareikalavus, visi asmenys ir įstaigos turi jam teikti valstybės statistikos reikalui juos liečiančių žinių.

Toms žinioms patikrinti Centralinis Statistikos Biuras turi teisę apžiūrėti atitinkamus dokumentus, bustus ir įrengimus.

Centraliniam Statistikos Biurui pareikalavus, viešosios teisės įstaigos, kurios renka statistines žinias, turi tų žinių jam pristatyti.

5.

Surinktos iš atskirų asmenų ir įstaigų statistikos žinios naudojamos tik statistikos reikalams. Jos gali būti paskelbtos arba išduotos tik kolektyvine arba nevardine forma.

6.

Centralinis Statistikos Biuras yra Finansų Ministerio žinioje. Jį veda direktorius.

7.

Svarbesniems statistikos klausimams svarstyti prie Centralinio Statistikos Biuro yra Statistikos Taryba.

8.

Statistikos Tarybą sudaro pirmininkas — Centralinio Statistikos Biuro direktorius — ir nariai: suinteresuotų ministerijų atstovai, po vieną atstovą nuo kiekvienos, Centralinio Statistikos Biuro redaktorius, Vytauto Didžiojo universiteto atstovas, Žemės Ūkio Rūmų atstovas, Prekybos ir Pramonės Rūmų atstovas ir Klaipėdos krašto Direktorijos atstovas.

9.

Statistikos Taryba svarsto šiuos dalykus:

1) Centralinio Statistikos Biuro metinį statistikos darbų planą;
2) viešosios teisės įstaigoms pavedamus statistikos darbus (3 str.);
3) bendruosius surašymo dėsnius ir visuotinių surašymų organizavimą;
4) naujų statistikos tyrinėjimo metodų įvedimą;
5) dėjimąsi prie statistikos konvencijų;
5) statistikos konferencijų šaukimą ir dalyvavimą tarptautinėse bei kitų valstybių šaukiamose statistikos konferencijose;
7) statistikos personalo ruošimą;
8) kitus dalykus, kuriuos jai paves svarstyti Finansų Ministeris.

10.

Statistikos Tarybos nutarimus tvirtina Finansų Ministeris, susitaręs su tuo ministeriu, kurio sritį tie nutarimai liečia.

11.

Privatinių įstaigų ir asmenų siunčiama Centraliniam Statistikos Biurui korespondencija su statistikos pranešimais atleidžiama nuo pašto rinkliavos. Tvarką šitai korespondencijai siųsti nustato Susisiekimo Ministeris, susitaręs su Finansų Ministeriu.

12.

Valstybės statistikos darbe dirbantiems asmenims ir įstaigoms centralinis statistikos biuras, Finansų Ministeriui sutikus, gali duoti pagyrimo ženklus arba atlyginimą pinigais.

13.

Šiam įstatymui vykdyti Finansų Ministeris leidžia taisyklių.

Baudžiamieji nuostatai.

14.

Asmuo, atsisakęs be teisėtos priežasties suteikti valstybės statistikos reikalui žinių arba suteikęs žinomai melagingų žinių, yra baudžiamas bauda ligi 250 litų.

21. Tarptautinės statistikos organizacija.

Tarptautinės statistikos darbo organizacija yra pradėta moderniškosios statistikos įkūrėjo Quetelet.

Tarptautiniai statistikos kongresai 1853—1876 m. Jo pasidarbavimu 1853 m. Briuselyj įvyko pirmasis tarptautinis statistikos kongresas. Tokių kongresų ligi 1876 m. įvyko net aštuoni. Juose dalyvaudavo atskirų valstybių oficialinių statistikos organų atstovai. Čia būdavo daugiausia svarstomi statistikos praktikos klausimai ir atskirų valstybių statistikos darbai mėginama suderinti. 1876 m. Budapešte šios rūšies kongresas buvo paskutinis. Mat, Vokietijos to laiko valdovas Bismarkas savo politikos tikslais Vokietijos statistikos atstovų daugiau nebeleido į tokius kongresus važinėti. O be Vokietijos daryti kongresus tais laikais kitos valstybės nedrįso ar tikslo nematė.

Vis dėlto Quetelet pradėtasis tarptautinės statistikos organizavimo darbas nežuvo. 1878 m. susirinko Paryžiuj dar 1872 m. Petrapilio statistikos kongreso numatytoji vad. nuolatinė tarptautinė statistikos komisija. Be to, tais pačiais 1878 m. Bertillon ir kt. Paryžiaus pasaulinės parodos metu organizavo tarptautinį demografijos kongresą, kuris vėliau posėdžiavo drauge su tarptautiniais higienos kongresais (ligi didžiojo karo tokių kongresų buvo 11, ir paskutinis 1912 Washingtone). Šituose kongresuose dalyvaudavo jau daugiau mokslo žmonės, kurie st. praktikos reikalų paprastai nesvarstydavo, o tenkindavosi demografijos metodų gvildenimu. Bet 1887 m. buvo įkurtas vad. Tarptautinis Statistikos Institutas.

Tarptautinis Statistikos Institutas (1887 m.). Jisai mėgino buv. Tarpt. Statistikos Kongresų darbą tęsti ir jį praplėsti. Būtent, Tarpt. Statistikos Instituto darbai pasireiškia ne tik oficialinių atskirų kraštų st. organų atstovų suvažiavimais, bet jo nuolatiniais dalyviais pakviečiami ir žymieji ar statistikos mokslui pasitarnavusieji atskiri mokslo žmonės. Tuo būdu Tarpt. Stat. Institutas statistikos mokslo problemų gvildenimus sujungė su st. praktikos klausimais. Be to, 1923 m. Haagoj buvo įkurtas nuolatinis Tarptautinis Statistikos Biuras (Office permanent), kuris savo leidiniuose renka ir skelbia visą jam prieinamą atskirų kraštų st. medžiagą.

Tarptautinio Statistikos Instituto suvažiavimai ligi karo būdavo daromi kas du metai. Karas tą visą darbą buvo sutrukdęs. Bet 1923 metais T. St. Instituto suvažiavimai atnaujinti. Šiame suvažiavime jau dalyvavo ir Lietuvos Centralinio Statistikos Biuro atstovas.

Kitokios tarptautinės statistikos įstaigos. Be to, prie Tarptautinės Telegrafo Sąjungos nuo 1864 m. veikia tarptautinės telegrafo statistikos skyrius ir nuo 1874 m. panašus skyrius prie tarptautinės pašto sąjungos (abidvi įstaigos Berne). Taip pat prie vad. Tarptautinio Darbo Biuro Bazelyj nuo 1901 m. pradėjo veikti tarptautinis darbo statistikos skyrius, kuris po karo drauge su visu T. D. Biuru teko Tautų Sąjungai Genevoj. Be to, 1905 m. Romoj įkurtasis Tarptautinis Žemės Ūkio Institutas turi taip pat savo gana gerai organizuotą žemės ūkio statistikos skyrių ir leidžia savo biuletenius. Taip pat yra st. skyrius, kuris leidžia mėnesinius biuletenius ir prie Tarptautinio Prekybos Rūmų Instituto Amsterdame.

Susikūrus Tautų Sąjungai, prie ekonominės sekcijos pradėjo veikti dabar jau plačiu mastu pastatytas tarpt. st. skyrius, kuris leidžia bendrojo pobūdžio mėnesinius statistikos biuletenius ir tarpt. st. metraščius, be to, specialinės socialekonominio gyvenimo klausimais (tarptautinė gamyba, rinka ir kainos, pinigai ir kreditas ir t. t.) leidžia dar ir statistinio pobūdžio monografijas.

Tarptautinės statistikos organizacijų pobūdį ligi tam tikro laipsnio turi ir vad. ūkio (ekonominės) konjunktūros institu-

tai. Nors šios rūšies įstaigų arba institutų svarbiausias uždavinys yra stebėti ir vaizduoti valstybės vidaus konjunktūros reiškinius, tačiau, dėl konjunktūros tarptautinių saitų, toki konjunktūros institutai negali išvengti (ir nevengia) ir pasaulinės ūkio konjunktūros reiškinių stebėjimo.

Savaime suprantama, kad šiais laikais kiekviena savarankiška valstybė, ar net mažiau ar daugiau autonomiški bei savivaldybiniai vienetai atskirose valstybėse, turi ir nuosavas statistikos įstaigas, arba statistikos biurus, kurių organizacija yra mažiau ar daugiau panaši į Lietuvos centralinio statistikos biuro ir jo darbų organizaciją.

22. Statistikos šaltiniai.

a) Statistikos šaltinių charakteristika.

Statistikos šaltinius tenka skirstyti į tris grupes: 1) pirminiai statistikos šaltiniai, 2) antraeiliai statistikos šaltiniai ir 3) statistinės kompiliacijos šaltiniai.

Pirminiai statistikos šaltiniai sudaromi iš statistikos duomenų, kurie yra gauti betarpiško (tiesioginio) arba tarpiško (netiesioginio) stebėjimo keliu. Kadangi tiek betarpiškieji, tiek tarpiškieji statistikos stebėjimai ligi šiol tedaromi tik valstybių arba autonominių vienetų ribose, tai ir pirminės statistikos šaltiniai tegali būti tik tautiniai arba valstybiniai, bet ne tarptautiniai. Tik kai kurios privatinio pobūdžio tarptautinės organizacijos, k. a. kai kurie tarptautiniai sindikatai, įpareigoja savo narius atskirose valstybėse betarpiškai teikti duomenis. Tokios organizacijos centro organas jau savo keliu tokius duomenis klasifikuoja ir paskelbia kaip tarptautinį statistikos šaltinį.

Pirminio statistikos šaltiniu pavyzdžiu Lietuvoje gali būti leidinys vardu „Lietuvos gyventojai“ arba kitas leidinys vardu „Visuotinis Lietuvos žemės ūkio surašymas“. Pirmasis leidinys yra sudarytas iš duomenų, gautųjų per 1923 m. visuotinį gyventojų surašymą, o antrasis — per 1930 m. visuotinį žemės ūkio surašymą. Taip pat pirminis st. šaltinis bus centralinio st. biuro kasmet leidžiamas leidinys vardu „Lietuvos užsienių prekyba“.

Antriniai arba antraeiliai (bet oficialiniai) statistikos šaltiniai bus mūsų Centralinio St. Biuro, taip pat kitų oficialinės statistikos įstaigų visi bendrojo pobūdžio iš įvairių gyvenimo sričių periodiniai ir neperiodiniai leidiniai, k. a.: mėnesiniai statistikos biuleteniai, statistikos metraščiai ir t. t. Antrinių statistikos šaltinių kategorijos priklauso ir beveik visi Tarptautinės Statistikos organų oficialiniai leidiniai (mėnesiniai biuleteniai, metraščiai ir kt.), kurie yra paremti betarpiškais tautinių ar atskirų valstybių statistikos organų praneštais st. duomenimis, arba tokiais pat ar panašiais duomenimis tarptautinių st. organų, kurie savarankiškai surenka juos iš pirminių arba kad ir antrinių, bet oficialinių, st. šaltinių. Šios kategorijos priklauso Tautų Sąjungos leidiniai: st. biuleteniai, st. metraščiai, taip pat Tarptautinio St. Instituto, Tarptautinio Žemės Ūkio instituto arba ir kitų kad ir tautinių, bet tarptautinio pobūdžio įstaigų, statistikos leidiniai.

Kompiliatyviniai st. šaltiniai dažniausiai yra sudaromi privatinių tyrinėtojų arba mažiau ar daugiau privatinio pobūdžio biurų ar įstaigų iš įvairių st. šaltinių. Tokios rūšies leidinį Lietuvoje, pavadintą „Statistinės žinios apie Lietuvą“, yra išleidęs ir vad. statistikos departamentas 1920 m. Šios kategorijos priklauso tarp kita ko ir Voitinskio leidinys „Die Welt in Zahlen“, taip pat Webb „The New Dictionary of Recent Statistics of the World to the Year 1911“, „The Statesman Yearbook, Statistical and Historical Annual of the States of the World“, Wappaeus, Stein ir Hoerschelmann „Handbuch der Geographie und Statistik“, Hübners „Geographisch - statistische Tabellen aller Laender der Erde“ ir kt.

b) Lietuvos statistikos šaltiniai.

I. Oficialiniai Centralinio Statistikos Biuro leidiniai:

1. Statistikos Biuletenis (leidžiamas kas mėnuo).
2. Lietuvos Statistikos Metraštis (leidžiamas kas metai).
3. Lietuvos Gyventojai. Kaunas, 1923, p. 311.

4. Lietuvos Apgyventos Vietos. Kaunas, 1925, p. 735.
5. Lietuvos Užsienio Prekyba (leidžiamas kas metai).
6. Visuotinis Lietuvos Žemės Ūkio Surašymas. 1930 m. I, pp. 164, II, pp. 58+188, II, pp. 72+92 ir IV, pp. 84.
7. Lietuvos Pasėlių Plotų ir Gyvulių Surašymas. Kaunas, 1935, p. 69.
8. Lietuvos Ūkis ir Rinka. Kaunas, 1932, p. 120.
9. Statistinės Žinios apie Lietuvą ligi Karui 1914 m. Kaunas, 1919, p. 200.
10. Lietuva Skaitmenimis 1918—1928 m. (diagramų leidinys).
11. Statistikos Žinios (Klaipėdos Krašto Statistikos Biuro hektografuotas leidinys).

II. Pusiau oficialiniai ir privatiniai statistikos šaltiniai.

12. Visa Lietuva (neperiodinis, kas 2—3 metai pasirodąs leidinys).
13. Kauno Miesto Statistikos Metraštis (pradėtas leisti 1936 m.).
14. Lietuvos Banko Biuletenis (eina neperiodiškai 4 kartus per metus nuo 1928 m.).
15. Rimka, A., Lietuvos Ūkis. Statistikos Tyrinėjimai. Kaunas — Vilnius, 1922, p. 135.

c) Bendrieji tarptautinės statistikos šaltiniai.

Tautų Sąjungos leidiniai (prancūzų ir anglų kalb.):

1. Bulletin Mensuel de Statistique (eina kas mėnuo nuo 1919 m.).
2. Annuaire Statistique (eina kas metai nuo 1927 m.).

II. Tarptautinio Žemės Ūkio Instituto (Romoje) leidiniai (prancūzų ir anglų kalb.):

3. Bulletin Mensuel de Statistique Agricole et Commerciale (eina kas mėnuo nuo 1909 m.).

4. Annuaire International de Statistique Agricole (eina kas metai nuo 1922 m.).

III. Tarptautinio Komercijos Instituto leidinys —

5. Recueil de Statistique de l'Institut International du Commerce (Briuselyje, eina kas mėnuo nuo 1921 m.).

IV. Tarptautinio Statistikos Instituto (Hagoje) leidinys —

6. Revue de l'Institut International de Statistique (eina kas 3 mėn. nuo 1933 m. su mėnesiniais kainų statistikos priedais [Supplement Mensuel]; prieš 1933 m. buvo leidžiamas Bulletin Mensuel maždaug to paties tipo, tik su siauresne programa, kaip Tautų Sąjungos statistikos biuletenis).

23. Statistikos vadovėliai ir periodinė literatūra.

a) Populiaresnieji st. teorijos vadovėliai.

(Žvaigždute * atžymėtosios knygos yra kuriuo nors atžvilgiu vertingesnės iš ta pačia kalba išleistųjų knygų).

Aftalion, A., Cours de Statistique, Paris, p. 247.
Ancyferov, A. N., Kurs statistiki. Praga, 1922, p. 228.
Bertillon, J., Cours Elementaire de Statistique administrative. Paris.
**Bojarsky, Brand* ir kt., Statistika. Moskva, 1936, p. 420.
**Connor, L. R.,* Statistics in Theory and Practice. London, 1932, p. 371.
Conrad-Hesse, I. Algemeine Statistik. Jena.
**Donner, Otto,* Statistik. Hamburg, p. 198.
**Harper, F. H.,* Elements of Practical Statistics. New York, 1930, p. 324.
G. von Mayer, Statistik und Gesellschaftslehre. I. Theoretische Statistik. Tübingen, 1914, p. 357.
**Julin, A.,* Principes de Statistique Theoretique et Appliquée. I. Statistique Theorique. Paris-Bruxelles.
Kaufmann, A., Teorija i Metody Statistiki.
Moncetz, A., Initiation aux Methodes de la Statistique. Paris, 1935, p. 80.
**Müller, J.,* Theorie und Technik der Statistik. Jena, 1927, p. 294.
Rimka, A., Statistikos Pagrindai. Kaunas, 1925, p. 190.
Schott, S., Statistik. Berlin, 1920, p. 128.

Wagenführ, R., Statistik leicht gemacht. Hamburg, p. 141.
Winkler, W., Statistik. Leipzig, 1933, p. 158.
Willford J. King., The Elements of Statistical Method. London.
Winkler, W., Grundriss der Statistik. Berlin, 1931, p. 290.
Žižek, F., Grundriss der Statistik. München und Leipzig, 1921, p. 470.

b) Matematinės statistikos vadovėliai ir vadovėliai su matematinio analizio metodų nuodugnesniu aiškinimu.

Anderson, O., Einführung in die Mathematische Statistik. Wien, 1935, p. 314.
Czuber-Burkhardt, Die Statistischen Forschungsmethoden. Wien, 1938, p. 330.
Fischer, A., The Mathematical Theory of Probabilities. New York, 1936, p. 289.
Yule, G. U., An Introduction to the Theory of Statistics. London, 1922, p. 415.
March, L., Les Principes de la Methode Statistique. Paris, 1930, p. 807.
Rietz-Bauer, Handbuch der Mathematischen Statistik. Leipzig und Berlin, 1930, p. 285.
Romanovsky, Elementarnyj kurs Matematičeskoj Statistiki. Moskva, p. 360.
Vinci, F., Manuale di Statistica. Bologna, 1936, p. 300.
Westergaard u. *Nybolle,* Grundzüge der Theorie der Statistik. Jena, 1928, p. 640.

c) Periodinė statistikos literatūra.

1) *Allgemeines Statistisches Archiv,* Jena. Eina kas trys mėnesiai. Vokietijos Statistikų Draugijos organas.
2) *Journal of the Royal Statistical Society,* London. Eina kas trys mėnesiai. Anglijos Statistikų Draugijos oficialinis organas.
3) *Journal of the American Statistical Association.* Eina kas trys mėnesiai. Amerikos Statistikų organas.
4) *Journal de le Société des Statistique,* Paris. Eina kas mėnuo.
5) *Revue de l'Institut International de Statistique, La Haye.* Eina kas 3 mėn. Rašiniai spausdinami prancūzų, anglų ir vokiečių kalbomis. Oficialinis Tarptautinio Statistikos Instituto leidinys.

ANTROJI DALIS.

A. Bendrieji statistikos teorijos ir metodologijos klausimai.

1. Statistikos metodų pagrindinės ypatybės.

Yra statistikų, kurie statistinį metodą tepripažįsta tik tais atvejais, kai statistinės datos arba vad. statistikos medžiaga yra gauta (arba gaunama) ištisiniu stebėjimu (visuotiniuoju surašymu). Šios G. v. Mayr'o absoliutinėj įtakoj palikusios statistikų grupės nuomonę statistinio metodo atžvilgiu labai ryškiai apibūdina J. Mueller'is, sakydamas, kad statistinis metodas pasižymi visuotinumu, arba, tiksliau — išsamumu (durch den Charakter des Erschoepfenden), ir šią tezę iliustruoja šiais pavyzdžiais:

„Statistinis metodas bus tuo atveju, kai gyventojų surašymo metu kiekvienas vietos asmuo su jo svarbiausiomis ypatybėmis į surašomąjį lakštą įrašomas; kai gimimai, vedybos ir mirimai pagal metrikacijos valdininkų užrašus be jokių praleidimų su įvairiausiomis ypatybėmis suskaitomi; jei, darant trobėsių surašymą, surašinėtojai eina nuo trobos prie trobos ir sužymi stebėjimui numatytus duomenis. O ne statistinis metodas bus tuo atveju, kai derėjimų vertinimuose ištiriami derliai nedaugelyj žemės sklypų ir tuomet daugybos keliu nustatomas bendrasis derlius didesniam plotui; jei kainos klausiamos nedaugelyj krautuvių ir šitokio stebėjimo (apklausinėjimo) duomenys laikomi charaketringais bendrajam kainų santykiavimui; taip pat nėra statistinis metodas, jei darbininkų gaunamo atlyginimo sąlygos yra nustatomos iš atlyginimo normų tūlose tarifinėse sutartyse ir kt.“[17].

Kitas šios statistikų grupės atstovas, prof. Fr. Žižek, statistinio metodo ypatybių ne ištisiniu stebėjimu paremtos sta-

[17] *Mueller,* Theorie u. Technik der Statistik. (Jena, 1929), p. 26.

tistikos griežtai neneigia; bet ir jis vadina tokią statistiką „anormaline“ arba „nekorektiška“. Tik Žižek jau pripažįsta, kad „nekorektiška“ statistika kai kada yra net geresnė už formališkai korektišką statistiką[17a].

O B o w l e y savo „socialinių reiškinių matavimuose“ tarp kita ko sako: „Jeigu iš surašymo mes žinome, kad duotojo (tam tikro) miesto 5% namų tenka daugiau kaip po 2 žmonės, jei mes įsitikinom, kad blogesnieji namai nėra dideli, jei mes kai kuriuos iš jų aplankėme, kad nustatytume faktinąsias žmonių gyvenimo sąlygas, jų amžių, lytį, verslą, — tai mes jau turime visus reikalingus duomenis, kad galėtume esamą padėtį kritikuoti arba siūlyti reformas, ir mums nereikės nei visų kambarių matuoti, nei visų be išimties namų lankyti[18].

Statistinio metodo pagrindinė ypatybė yra ta, kad, visų pirma, statistinis stebėjimas apima m a s ę, ir, antra, statistinis vaizdavimas turi tikslo parodyti masės c h a r a k t e r į, t. y. pačias pagrindines arba esmines ypatybes. Kuriuo būdu tasai masės stebėjimas daromas, arba kuriomis statistinės technikos priemonėmis tos masės charakteringųjų ypatybių vaizdavimas atliekamas — tai antraeilis klausimas. O iš to seka, kad statistinio metodo charakterį nulemia ne masės stebimųjų vienetų absoliutinis pilnumas arba išsamumas, o aplinkybės, ar pasirinktas statistinio stebėjimo metodas įgalina stebimos masės charakterį, t. y. jos pagrindines ypatybes, pažinti ir atvaizduoti, ar neįgalina.

Šita pagrindinė statistinio metodo ypatybė savaime nulemia ir statistinio metodo aiškiai apibrėžtą santykį su didžiųjų skaičių dėsniu. Ir tobuliausiai paruoštas bei atliktas ištisinis stebėjimas, arba visuotinis surašymas, nieko bendra su statistiniu stebėjimo metodu neturės, jei bus stebimi ir žymimi toki skaitomieji vienetai, kurie atstovaus tik patys save ir nebus jiems panašių nei artimų vienetų didesnio skaičiaus. Tokio ištisinio ir „be jokių praleidimų“ daromo stebėjimo pavyzdžiu yra bankų ir kai kurių verslo įmonių arba namų ūkio sąskaityba.

[17a] *Žižek,* Grundriss der Statistik (Muenchen, 1921), 187 ir palyg. to paties autoriaus „Nichtkorrekte statistische Verfahren“, atsp. Allg. Statistisches Archiv, 21, Bd., I, Heft (1931), p. 52.

[18] Cituota iš rusų kalbos vertimo, pavadinto „Osnovy socialnoj statistiki“, Moskva, 192, pusl. 5/6.

Šiais atvejais kiekvienas turto arba inventoriaus objektas yra labai stropiai ir atsidėjus stebimas ir su visomis ypatybėmis atžymimas arba įrašomas; taip pat ir kiekvienas reiškinys visai tiksliai be jokių praleidimų registruojamas, vienodi arba giminingi reiškiniai grupuojami ir aiškiai aprėžtais skaitomaisiais periodais sumuojami lygiai taip, kaip sumuojami skaitomaisiais periodais užsienių prekybos arba natūrinio gyventojų judėjimo duomenys. Vis dėlto bet kurio ūkio, arba keletos ūkių, sąskaitybos reikalui daromas „be jokių praleidimų“ stebėjimas bei žymėjimas nieko bendra neturi su statistika nei su statistiniu metodu, — nors savaime visai tokio pat arba net ir to paties stebėjimo duomenys gali būti statistikos naudojami. Tačiau pirmoji ir pagrindinė sąlyga, kuri turi būti pildoma norint sąskaitybos ar bet kurio kito „be praleidimų“ stebėjimo duomenis statistikai panaudoti, yra ta, kad tie duomenys būtų didžiųjų skaičių dėsnio veikiami. O tai reiškia, kad tokių duomenų turi būti tiek, kad jais jau būtų galima atvaizduoti masę arba masės charakterį. Ar šita sąlyga bus išpildyta stebimųjų vienetų skaičių plečiant ligi to, kad būtų galima sakyti, jog stebima yra jau „masė“, ar tik skaitomųjų invervalų ar periodų skaičių plečiant, kad skaitomieji vienetai būtų veikiami didžiųjų skaičių dėsnių laiko vienetų masėje — esminio skirtumo nesudarys. Svarbu tiktai, kad didžiųjų skaičių dėsnis dalykų stovio arba reiškinių laiko skaitomųjų vienetų masėje pasireikšti galėtų.

Iš to seka ir kita statistinio metodo ypatybė: ir su „praleidimais“ arba „ne pilnas“ st. stebėjimas bei žymėjimas atitiks statistikos metodo būtinas sąlygas arba reikalavimus, jei tik bus stebimi ir žymimi tikrieji masės reprezentantai didžiųjų skaičių dėsningo veikimo (arba pasireiškimo) aplinkybėse, t. y.: arba tie reprezentantai bus „tipiški“ ir masės daugumui charakteringi (atstovaus masės daugumą) arba — patys reprezentantai bus atsitiktiniai ir masę pilnai galės atstovauti kaip tik dėl savo atsiktinumo, kuris jau buvo nulemtas didžiųjų skaičių dėsningo veikimo. Tuo būdu statistinis metodas betarpiškai yra susijęs iš vienos pusės su didžiųjų skaičių dėsningu veikimu, iš kitos — su pripuolamumu arba atsitiktinumu.

2. Kai kurios statistinės sąvokos ir terminai.

Jau įvaduose buvo pabrėžta, kad statistika yra masių arba masinių reiškinių žymėjimas skaitmenimis ir jų vaizdavimas charakteringomis ypatybėmis. Vadinasi, visų pirma, statistikai yra reikalinga masė, kurią gali sudaryti daiktai ir dalykai, arba masiniai reiškiniai, t. y. masiniai įvykiai bei atsitikimai. Nepriklausomai nuo to, ar mums rūpimą masę sudaro daiktai ar kai kurie reiškiniai, sutarsime ją vadinti statistine mase.

Statistinė masė gali būti statistikos metodais stebima, vaizduojama ir analizuojama.

Statistinis stebėjimas gali vykti statistinės masės vienetų, arba tų vienetų kolektyvų (grupių) surašinėjimu arba registracija. Šiam reikalui tenka išskirti ir aiškiai apibrėžti statistinės masės skaitomuosius (stebimuosius) ir surašomuosius vienetus.

Skaitomasis (stebimasis) vienetas statistinėj masėj bus kiekvienas daiktas, dalykas, individas, reiškinys, atsitikimas ar kt., kuris vienu ar kitu statistikos metodu stebimas ir surašomajame lakšte skaitmeniniu vienetu pažymimas.

Surašomasis vienetas gali su skaitomuoju vienetu sutapti, bet gali ir nesutapti: o sutaps surašomasis vienetas su skaitomuoju vienetu, kai, darant statistinį stebėjimą, kiekvieno atskirai paimto skaitomojo vieneto ypatybės bus atžymimos atskirai; nesutaps surašomasis vienetas su skaitomuoju, kai skaitomieji vienetai, juos stebint, jau iš karto jungiami į atskiras grupes (kolektyvus) pagal aiškiai apibrėžtas ypatybes. Pav., sutaps skaitomasis vienetas su surašomuoju vienetu, kai, darant gyventojų surašymą, kiekvieno asmens ypatybės (lytis, tautybė, amžius ir kt.) bus žymimos paskirai. Nesutaps — kai iš karto bus atžymėta tik bendrieji kurią nors vieną ypatybę turintys vienetų skaičiai (tiek arba tiek vyrų, moterų, tiek arba tiek vaikų, suaugusių ir t. t.).

Statistikos metodais stebimos arba vaizduojamos skaitomųjų vienetų ypatybės yra vadinamos statistinėmis žymėmis ir pažymiais.

Surašomoji st. žymė yra vadinama tokia st. skaitomojo vieneto ypatybė, kurią stebimos masės vienetai turi, nors pažymiai gali būti ir nevienodi. Pav., kiekvienas žmogus priklauso kokios nors tautybės, moka kurią nors kalbą, turi vienokį ar kitokį verslą ar šiaip bet kurį gyvenimo šaltinį: bet ne kiekvienas žmogus yra tos pačios tautybės, ne visi žmonės tą pačią kalbą moka, ne visi žmonės iš to paties verslo gyvena. Taigi darant gyventojų surašymą kiekvieno iš jų tautybės surašomoji žymė turės būti atžymėta kuriuo nors pažymiu: lietuvis, lenkas, vokietis, rusas ar kt. Amžiaus (surašomoji) žymė turės būti atžymėta metų pažymiu. Pilietybės žymė — atitinkančios valstybės pažymiu, ir t. t.

Statistinė eilutė. Kai statistinė masė vienomis ar kitomis žymėmis arba jų pažymiais yra suskaidyta ir skaičiais atžymėta, tat tokia masė jau sudaro skaičių eilutę, kuri šiuo atveju vadinama statistine eilute. Vadinasi, ne kiekviena skaičių eilutė yra statistinė eilutė, o tik tokia, kuri yra kilusi iš statistinės masės sudarytųjų duomenų.

Statistinės masės stebimieji duomenys yra reikalingi atžymėjimo tam tikrame surašomajame lakšte.

Surašomasis lakštas statistikos praktikoj turi ne vieną vardą. Kai kada jis vadinamas surašomąja kortele, registracijos kortele, statistikos deklaracija ar kt.

Sulig tuo, ar į surašomąjį lakštą atžymimas kiekvienas skaitomasis vienetas atskirai, ar į vieną tokį lakštą atžymimos ypatybės viso kolektyvo, surašomieji statistikos lakštai vadinami individualiniais arba kolektyviniais (bendriniais). Ta pačia prasme kalbama ir apie individualinį ar kolektyvinį (bendrinį) st. stebėjimą.

Skaitomoji (stebimoji) diena. Kai st. metodais stebimi kurie nors daiktai jų skaičiui ir sudėčiai bei padėčiai nustatyti, yra reikalinga iš anksto paskirti tam tikrą vieną dieną, kurią stebimos masės (daiktų) stovis (padėtis) surašomajame lakšte atžymimas. Tokia diena vadinama skaitomoji arba stebimoji diena.

Surašomoji diena. Patsai surašymas gali vykti tą pačią dieną, bet gali vykti ir kitą dieną arba užsitęsti ir ilgesnį

laiką. Diena, kurią atliekamas pats surašinėjimas, t. y. skaitomojo vieneto žymių ir pažymių į surašomąjį lakštą įrašymas, vadinama surašomoji diena.

Skaitomasis periodas. Jei statistinis stebėjimas liečia ne daiktus ar jų stovį, o kuriuos nors reiškinius ir jų judėjimą, tai skaitomąją dieną atstoja vadinamas skaitomasis periodas.

Kritiškasis momentas. Tiek stovio, tiek judėjimo statistikoj, be aiškiai aprėžtos skaitomosios dienos (arba skaitomojo periodo), yra dar reikalinga griežtai aprėžti tam tikras momentas, kuriam visos skaitomojo vieneto surašomos žymės fiksuojamos (į surašomąjį lakštą įrašomos). Šis momentas vadinamas kritiškuoju momentu.

Statistikos stebėjimo (surašymo) sistemos. Į vieną surašomąjį lakštą gali būti įrašoma vieno skaitomojo vieneto surašomos žymės su pažymiais, bet gali būti į vieną surašomąjį lakštą surašoma ir keleto skaitomųjų ir surašomųjų vienetų žymės su pažymiais. Pirmuoju atveju kiekvienas skaitomasis vienetas turi savo atskirą kortelę; antruoju — gaunamas visas skaitomųjų ar surašomųjų vienetų sąrašas. Taigi patsai surašymas arba statistinis stebėjimas vykdomas arba kortelių sistema arba sąrašų sistema.

3. Statistikos darbo etapai ir jų uždaviniai.

Yra trys statistikos darbo etapai: a) st. stebėjimas, b) st. vaizdavimas ir c) st. analizis. Šie trys statistinio darbo etapai yra betarpiškai susieti. Tačiau gali būti st. duomenys jau surinkti, vadinasi, patsai statistinis sebėjimas palieka nebereikalingas. Gali taip pat būti st. duomenys ne tik surinkti, bet ir mažiau ar daugiau suklasifikuoti, t. y. ligi tam tikro laipsnio atliktas ir statistinio vaizdavimo darbas. Tokiu atveju iš karto galima eiti prie st. duomenų analizio, kurio priklauso ir aukštesnės kategorijos statistinis vaizdavimas, t. y. statistinė klasifikacija.

Tačiau ir tais atvejais, kai betarpiškas st. stebėjimas nebereikalingas, arba kai st. duomenys yra sugrupuoti ir paskelbti st. šaltiniuose, vis dėlto prasmingas st. duomenų analizis nėra galimas, jei nėra žinomi metodai, kuriais st. stebėjimas buvo

darytas, nei principai, kuriais pirminis st. duomenų vaizdavimas padarytas. Taigi, ir st. duomenų vartotojui arba naudotojui yra reikalinga pažinti ne tik st. šaltinius, bet ir statistinio stebėjimo metodus bei st. vaizdavimo principus.

Pirmasis st. darbo uždavinys — visapusiškai pažinti stebimos (vaizduojamos, analizuojamos) masės arba masinio reiškinio ypatybes, kad būtų galima st. metodais atvaizduoti tos masės ar reiškinio pačius charakteringuosius bruožus. Vadinasi, statistika visuomet turi reikalą su mase ar masėmis, arba masiniais reiškiniais bei atsitikimais. O kad būtų galima prasmingai atvaizduoti masės charakteringus bruožus, yra reikalinga imti tą masę tokią, kokia ji yra n o r m a l i o s e sąlygose, t. y. normalioje padėtyje bei sudėtyje.

Tuo būdu prasmingas masinių reiškinių arba atsitikimų stebėjimas bei vaizdavimas yra reikalingas nuolatinio arba sistemingo tų reiškinių arba atsitikimų eigos (kitimų) stebėjimo ir jų žymėjimo. Šia prasme kalbama apie s t o v i o arba p a d ė t i e s statistiką ir apie r e i š k i n i ų arba a t s i t i k i m ų statistiką. Statistinio stebėjimo sąlygos vienu ir antru atveju yra kiek skirtingos.

B. Statistinio stebėjimo pagrindinės sąlygos.

4. Masių stovio arba padėties statistinis stebėjimas.

Jei norime statistinio stebėjimo keliu nustatyti arba surasti gyventojų, ūkių ar įmonių, gyvulių, namų ar kitų kokių daiktų skaičių, jų sudėtį bei padėtį, tai pirmiausia reikia parinkti tokiam stebėjimui atitinkamą l a i k ą ir momentą bei žiūrėti, kad ir kitos a p l i n k y b ė s būtų palankios. O atitinkamo momento aplinkybes stovio statistikai sudaro trys dalykai: 1) skaitomoji diena, 2) kritiškasis momentas ir 3) surašomoji diena.

a) S k a i t o m o j i d i e n a s t o v i o s t a t i s t i n i a m s t e b ė j i m u i negali būti parenkama savivalingai, jei norima, kad statistinis stebėjimas įgalintų atvaizduoti charakteringas stebimos masės ypatybes. Pav., jei norima per gyventojų surašymą sužinoti ne tik gyventojų skaičių, bet ir jų pasiskirsty-

mą atskirose vietovėse (kiek žmonių gyvena miestuose ir kiek sodžiuose), tai būtų maža prasmės mūsų sąlygose tokį surašymą daryti, pav., liepos mėn., kai miestų gyventojų dalis yra išvažinėjusi į sodžių ar kurortuose atostogauja. Taip pat ir gyventojų verslo duomenys gali būti iškreipti, jei statistinis verslo stebėjimas būtų daromas tuo metu, kai dėl vienų ar kitų priežasčių atskirose ūkio šakose yra susikaupę sezoniniai darbai (arba sezoninis nedarbas). Gyvulių statistinis stebėjimas tais pačiais metais duos visai skirtingus skaičius, kurie pareis vien nuo skaitomosios dienos parinkimo: skaitomąja diena parinkus kurią nors dieną pavasarį, kai yra daug jauniklių, — gyvulių skaičius išeis daug didesnis kaip vėlai rudenį, kada tokių jauniklių maža ir, aplamai, gyvulių perteklius būna likviduotas, ar palikę tik vadinami „žieminiai" gyvuliai.

Skaitomoji diena stovio statistiniam stebėjimui turi būti parenkama tokia, kad stebimoji masė (gyventojai, gyvuliai, ūkiai ar kt.) būtų kiek galint normalesniame stovyje visais atžvilgiais. Be to, kad išvengti galimų praleidimų ir dvigubų skaičiavimų, negana nustatyti skaitomąją dieną, o reikalinga gana griežtai apibrėžti ir vad. stebėjimo kritiškąjį momentą.

b) Stebėjimo kritiškasis momentas taip pat reikia parinkti atsiželgiant pagrindinio statistikos reikalavimo (kad stebimoji masė būtų kiek galint normalesnėse sąlygose). Pav., darant gyventojų st. stebėjimą, surašomuoju vietovės pažymiu gali būti kiekvieno asmens tam tikroj vietoj kritiškuoju stebėjimo momentu buvimo faktas arba nuolatinis gyvenimo pažymys. Jeigu skaitomąja diena būtų parinktas šventadienis su kritiškuoju momentu 12 val. dieną, kai sodžiaus gyventojų dauguma būna bažnyčioj, o vietovės pažymys būtų buvimo faktas, tai gautieji statistikos duomenys rodytų perpildytus miestus, miestelius ir bažnytkaimius — ir beveik tuščius sodžius bei kaimus. Dėl to ir 1923 m. darant visuotinį gyventojų surašymą Lietuvoje kritiškasis stebėjimo momentas buvo parinktas rugsėjo (11) 17 d. ne 12 val. dienos, o 12 val. nakties. Tačiau jeigu 12 val. nakties kritiškam momentui būtų nustatytas, pav., ir Londono miesto gyventojams, tai patsai Londono centras, vadinamas City (kur dieną gyvena ir dirba žmonės susispaudę kaip silkės), išeitų beveik negyvenamas, — ka-

dangi naktį Londono miesto centre tegyvena tik vieni namų sargai.

Šie pavyzdžiai rodo, kad kritiškasis statistinio stebėjimo momentas turi būti pritaikintas ne tik kiekvienos stebimos masės arba objekto charakteringoms ypatybėms apskritai imant, bet turi būti atsižvelgiama ir įvairių vietos bei kitų aplinkybių, kurios tą charakterį gali vienon ar kiton pusėn pakreipti.

Tikslingas kritiškojo momento parinkimas yra svarbus dar ir kitais atžvilgiais. Pav., jei darant statistinį stebėjimą būtų atsižvelgiama vien tik skaitomosios dienos be griežtai aprėžto momento, tai niekas negalėtų susivokti, kaip pasielgti tais atvejais, kai skaitomąją dieną stebimoj masėj įvyksta vienos ar kitos rūšies pakitimai, k. a. kas nors numiršta, gema, išvyksta, atvyksta ir t. t. Nustačius gi kritiškąjį momentą, sakysim, 12 val. naktį (arba kitą kurį momentą), statistikos stebėjimas (surašymas) liečia tik tuo stebimus vienetus (gyventojus, gyvulius, namus ar kt.), kurie tuo griežtai aprėžtu momentu buvo ar tam tikras žymes turėjo. Ir taip, pav., 1923 m. rugsėjo 18 d. ir kt. dienas surašinėtojai tik tuos gyventojus surašė, kurie rugsėjo 17 d. vidurnaktį gyveno. Jei kas nors pasimirė (ar išsikėlė) bent vieną minutę prieš vidurnaktį, jis nebuvo surašomas, — nežiūrint į tai, kad visą dieną dar gyvas ir sveikas buvo arba toj vietoj gyveno. O jeigu kas gimė ar į tam tikrą vietą atsikėlė bent vieną minutę po vidurnakčio, jis taip pat negalėjo būti įrašytas, — nežiūrint į tai, kad surašomąją dieną (rugsėjo 18 ar kt.) jis jau toj vietoj gyveno. Bet užtat buvo įrašyti visi tie, kurie gimė arba atsikėlė į kurią vietą skaitomąją dieną dar prieš vidurnaktį, — nežiūrint į tai, kad gimimo (arba atsikėlimo) momentas ir vad. kritiškasis stebėjimo momentas būtų nesutapę.

Lygiai taip pat elgiamasi statistikos metodais stebint ne žmones, o gyvulius, namus, inventorių ar šiaip bet kurios rūšies daiktus arba gėrybes.

c) Surašomoji diena arba surašomasis laikas statistiniam stebėjimui tokios didelės reikšmės, kaip skaitomoji diena ir kritiškasis stebėjimo momentas, neturi. Pav., per visuotinį Lietuvos gyventojų surašymą 1923 m. skaitomoji diena buvo nustatyta rugsėjo 17 d., bet surašinėjama bu-

vo keletą dienų (ligi rugsėjo 25 d.) Per visuotinį žemės ūkių surašymą 1930 m. skaitomoji diena buvo gruodžio 30 d., o surašomasis laikas tęsėsi ligi sausio 10 d. 1931 metų.

Vis dėlto nėra pageidaujama, kad surašomasis laikas perdaug nusitęstų. Priežastis ta, kad juo ilgesnis laikas bus praėjęs nuo skaitomosios dienos kritiškojo momento ligi surašomosios dienos, tuo tenka laukti didesnių paklaidų stebimuose duomenyse. Idealiausias surašomasis laikas bus tas, kuris bus kritiškajam momentui arčiausias. Dar nedaug paklaidų teįvyks, jei statistinis stebėjimas (surašymas) bus atliktas per keliolika valandų po kritiškojo momento. Ir juo surašymas ilgiau užsitęsia, tuo paklaidų turi būti daugiau: nes ne visų klausiamųjų asmenų atmintis yra lygiai gera.

5. Judėjimo arba reiškinių bei atsitikimų statistinis stebėjimas.

Jei norime statistinio stebėjimo kelių nustatyti, pav., gyventojų natūrinį arba vad. mechaninį judėjimą (gimimus, mirimus, imigraciją), tai stebėjimo sąlygos jau bus kitokios. Skaitomoji diena su vienu kritiškuoju momentu mums patiektų tik to momento mirimus, gimimus ar kt. O kiek iš viso buvo tokių atsitikimų kad ir per ištisą dieną — mes patirti negalėsime. Taip pat toli gražu nepilnus duomenis gautume ir surašomąjį laiką aprėžę. O reiškinių bei atsitikimų statistikos uždavinys — parodyti tų reiškinių arba atsitikimų j u d ė j i m o charakteringas ypatybes, t. y., kaip tie reiškiniai arba atsitikimai k i n t a laikui bėgant. Dėl to skaitomąją dieną judėjimo statistikoj turi pakeisti s k a i t o m a s i s p e r i o d a s, kuris a p r ė ž i a m a s d v i e j ų k r i t i š k ų j ų m o m e n t ų. Patsai gi statistinis stebėjimas, arba surašomasis laikas, gali būti aprėžtas tik atskirais atvejais praeities atsitikimams suregistruoti. Aplamai, toksai stebėjimas daromas n u o l a t o s, t. y., įvykiai arba atsitikimai surašinėjami nuolatos ir kiek galint greičiau po to, kaip tik jie pasireiškia arba būna pastebėti.

a) S k a i t o m a s i s p e r i o d a s judėjimo statistikoj taip pat turi būti parinktas toks, kad įgalintų stebimų reiškinių ar atsitikimų charakteringas ypatybes teisingai pastebėti ir atvaizduoti. Pav., gyventojų natūriniam judėjimui (gimimams, mi-

rimams, vedyboms) stebėti skaitomuosius periodus galima parinkti metinius, mėnesius arba kitokius — pagal reikalą ir aplinkybes: nes šios rūšies įvykiai, apskritai imant, su nė per dideliais svyravimais atsitinka nuolatos per metų metus. Tačiau užsienių prekybos statistikai vien metinių skaitymo periodų nustatymas neduotų galimumo pastebėti visos eilės sezoninių prekių prekybos eigos. O javų derliams, gyvulių natūriniam prieaugliui statistikos metodais stebėti tegali prasmės turėti tik metiniai skaitymo periodai, — kadangi pasikartojimai ir reikšmingesni pasikeitimai, aplamai, įvyksta vieną kartą per metus. Iš kitos pusės, oro temperatūros statistiniam stebėjimui jau bus perilgi ne tik metiniai skaitymo periodai, bet ir mėnesiniai, — kadangi būna laiko tarpai, kai oro temperatūros skirtumai atskirų mėnesių pradžioj ir gale būna labai dideli (pav., gruodžio arba kovo mėn., taip pat balandžio ir spalių mėn.). Dėl šios priežasties oro temperatūros ir, aplamai, klimato reiškinių statistikoj skaitomieji periodai aprėžiami dažniausiai dekadomis (kas 10 d.), ir kiekvienas mėnuo dalomas į tris tokias dekadas.

Tačiau jau pastarasis pavyzdys pirštu prikišamai parodo, kad skaitomojo periodo aprėžimas judėjimo statistikai susitinka su daug didesnėmis sunkenybėmis, kaip skaitomosios dienos nustatymas stovio statistikai. O priežastis ta, kad skaitomiems periodams skaičiuoti mes neturime griežtai aprėžto ir visiems atvejams tinkamo vienodo masto. Pav., 1 dienos skaitomuosius periodus mes be jokių sunkenybių galime grupuoti į savaitinius periodus, arba atvirkščiai — savaitinius galime skaidyti į dienos periodus. Tačiau savaitinių periodų mes negalime jungti į mėnesinius periodus, nes iš 12 mėn. tik vasaris (ir tai ne visuomet) turi lygiai 4 savaites. Taip pat ir mėnesiai: vieni turi 30 dienų, kiti 31 d., vasaris gi tris metus iš eilės turi 28 dienas, o ketvirtaisiais metais 29 d., — vėl išskiriant 100-uosius metus, kai vasaris paliekamas tik su 28 dienomis. Taip pat ir metai turi arba 365 arba 366 dienas.

Ši aplinkybė ne tik apsunkina vienodu tikslumu griežtai apibrėžti skaitomąjį periodą judėjimo statistikai ir paskui trumpesnius periodus jungti į ilgesnius (arba ilgesnius skaidyti į trumpesnius), bet mūsų vartojamų laikui matuoti skaitomųjų periodų nevienodumas sunkina ir statistinio vaizdavimo bei analizio darbą. Pav., jei mėnesiai būtų vienodo ilgumo, tat

bet kuris skirtumas judėjimo duomenyse būtų kitimų rodikliu. Bet kai vieni mėnesiai turi 30 dienų, kiti 31 d., o dar kiti 28 arba 29 d., tai mėnesinių duomenų nedideli skirtumai savaime dar nieko nepasako: nes jie gali pareiti ne nuo stebimo reiškinio judėjimo kitimų, o vien nuo skaitomojo periodo skirtumų.

b) K r i t i š k a s i s m o m e n t a s judėjimo statistikoj turi būti griežtai nustatytas kiekvieno skaitomojo periodo pradžiai ir galui. Tik tie reiškiniai arba atsitikimai kiekvienam skaitomajam periodui žymėtini, kurie įvyksta anksčiau buvusiam skaitomajam periodui nepasibaigus ir vėliau einančiam periodui dar neprasidėjus.

Jei skaitomasis periodas tęsiasi ištisą dieną, kiekvieno periodo kritiškieji momentai gali būti aprėžti vidurnakčiu. Bet jeigu stebimųjų atsitikimų dauguma paprastai vyktų ne dienomis, o naktimis, tat vidurnakčio parinkimas kritiškuoju momentu skirtų naturalius skaitomuosius periodus ir padarytų iš jų nenaturalius, grynai dirbtinius, ar net galėtų iškreipti stebimo reiškinių judėjimo charakteringas ypatybes. Taip pat ir metinių skaitomųjų periodų aprėžimas vad. kalendoriniais metais gali iškreipti tuos judėjimo statistikos duomenis, kurie turi savo naturalius judėjimo periodus (arba sezonus) kitokius, negu vad. kalendoriniai metai. Pavyzdžiu gali būti žemės ūkio produktų eksporto statistinis stebėjimas. Jei, sakysim, iš vienų metų derliaus kurio nors javo buvo eksportuota 5000 tonų sausio — birželio mėnesiais ir iš kitų metų derliaus 5000 tonų rugsėjo — gruodžio mėnesiais, tai tų metų eksporto statistikoj būtų parodyta to produkto buvus eksportuota 10.000 tonų, — nežiūrint į tai, kad toksai kiekis tebuvo eksportuotas tik iš dviejų metų derliaus perteklių. Ir atvirkščiai: kitam skaitomajam periodui gali netekti nieko arba labai mažai, jei dėl atsitiktinių eksporto sezono nukrypimų kitų metų eksporto skaitomajam periodui būtų tekę dviejų metų derliaus pertekliai.

Šis pavyzdys rodo, kad ir judėjimo statistikoj kritiškojo momento nustatymas skaitomajam periodui turi būti daromas atsižvelgiant stebimojo reiškinio charakteringų ypatybių, šiuo atveju — atsižvelgiant stebimo reiškinio daugumos susikaupimų (dispersijos) laiko bėgyje. Jei kurios nors rūšies reiškiniai arba atsitikimai vyksta nuolatos, be charakteringų (žymesnių)

susikaupimų atskirais periodo laikotarpiais, tai skaitomųjų periodų ir jiems kritiškųjų momentų nustatymas gali būti daromas sulig nuožiūra, arba pagal pasistatytą tikslą bei reikalą. Bet jeigu stebimas reiškinys turi charakteringų judėjimo ypatybių (susikaupimų), arba jeigu tas judėjimas pareina nuo sezoninių priežasčių, tai skaitomasis periodas turi būti apibrėžtas taip, kad jis atitiktų ir natūrinio judėjimo susikaupimų periodus. Kritiškasis gi momentas turi būti nustatytas taip, kad jis atitiktų kiek galint mažesnio judėjimo laikotarpį. Čia tenka elgtis lygiai taip, kaip elgiamasi parenkant kritiškąjį momentą stovio statistiniam stebėjimui (parenkamas iš to laiko, kai stebimoji masė būna „ramiausia", pav., gyventojams vidurnaktis ir pan.). Dėl šios priežasties javų derliams stebimasis periodas gali būti imamas ir vad. kalendorinių metų 12 mėnesių laikas; tačiau, javų derliaus eksportui kalendorinius metus paėmus, visuomet pateks dviejų metų derliaus eksporto duomenys. Tokiu atveju bus tiksliau metinį periodą aprėžti ne kalendoriniais metais, o vadinamais „ūkio metais", kurie, paprastai, skaitomi arba nuo birželio mėn. 30 d. ligi sekančių kalendorinių metų liepos 1 d., arba nuo liepos 31 d. ligi rugpiūčio 1 d.

c) Stebimasis arba surašomasis laikas reiškinių arba atsitikimų judėjimo statistikoj gali būti iš anksto aprėžtas arba ir nuolatinis. Tačiau ir judėjimo statistikoj palankiausios stebėjimo sąlygos bus tuo atveju, kai stebėjimo laikas bus kiek galint artimesnis stebimam periodui. Dėl šios priežasties, pav., natūrinio gyventojų judėjimo, užsienių prekybos ir kt. statistinis stebėjimas (medžiagos rinkimas) daromas nuolatinės registracijos būdu, o duomenys kiekvieno skaitomojo periodo gale tik susumuojami. Nuolatinis reiškinių arba atsitikimų judėjimo stebėjimas dar ir tuo geresnis, kad jis įgalina kiekvienu metu turėti pačius naujausius statistikos duomenis. Nenuolatinis, arba laiko atžvilgiu aprėžtas, reiškinių arba atsitikimų stebėjimas tegali būti tikslingas tik netiesioginiu keliu renkant medžiagą iš susikaupusių duomenų kurioj nors ne statistikos įstaigoj, arba kai patys atsitikimai vyksta ne nuolatos, o pasikartoja tam tikrais intervalais. Pav., metinio derliaus statistinį stebėjimą gali būti tikslinga aprėžti javų valymo periodu,

sezoninių prekių apyvartos statistinį stebėjimą—atitinkamų sezonų periodu, ir t. t. Taip pat dar gali būti tikslinga aprėžti statistinį stebėjimą tokių reiškinių, kuriems taikomas reprezentacinis metodas. Pav., pieno ūky gautasis pienas, arba jo produktai, tiksliai stebimi tik vieną dieną per savaitę ar net per mėnesį; prekių kainos stebimos tik vieną — antrą dieną per mėnesį, ir t. t.

Tuo būdu stebimojo arba surašomojo laiko apibrėžimas bei nustatymas reiškinių arba atsitikimų judėjimo statistikoj turi būti daromas atsižvelgiant stebimųjų reiškinių pobūdžio ir jų pasireiškimo aplinkybių, be to, reikalinga atsižvelgti ir statistinio stebėjimo uždavinių bei metodų.

C. Statistinio stebėjimo metodai ir sistemos.

6. Bendrosios pastabos.

Kiekviena masė arba masinis reiškinys gali būti stebima įvairiais būdais. Pav., gyventojų skaičiui bei sudėčiai surasti gali būti daromas visuotinis gyventojų surašymas. Bet savo keliu gyventojai, paprastai, dar yra registruojami policijos arba savivaldybių įstaigose; taigi statistikai reikalingi duomenys gali būti imami ir iš tų įstaigų. Natūrinis gyventojų judėjimas pas mus ir kitur šiandien nuolat registruojamas vad. civilinių aktų sudarymo įstaigose, ir tos registracijos rezultatai betarpiškai siunčiami statistikos organams. Tačiau tuos pačius duomenis, kad ir gal ne su visais pažymiais, galima rasti ir tų pačių įstaigų knygose arba aktuose, be specialinio statistikos reikalui kortelių surašinėjimo ir be jų siuntinėjimo Centr. St. Biurui. Taip pat ir užsienių prekybos duomenis (kad ir ne visai pilnus) apie prekių kiekį ir rūšį galima rasti muitinėse ar jų archyvuose net ir tuo atveju, jei vadinamos statistikos deklaracijos nebūtų specialiai rašomos.

Iš kitos pusės, apie gyventojų kiekį ir jų padėtį galima gauti duomenų stebint ir žymint kiekvieną asmenį atskirai, — arba stebint visos šeimos (sodžiaus, ūkio, valsčiaus) gyventojus kartu, suskirsčius juos į grupes tam tikrais iš anksto nustatytais pažymiais. Taip pat skirtingas stebėjimo ir žymėjimo

metodas gali būti pritaikintas gyvulių, namų, gimimų, mirimų ir kt. statistikai.

Be to, statistikos praktikoj vieni dalykai arba reiškiniai stebimi ir žymimi visi be jokios išimties, o kitais atvejais būna daromas tik dalinis stebėjimas bei žymėjimas pagal kurią nors pasirinktą sistemą.

Tuo būdu statistinio stebėjimo metodų bei sistemų yra įvairių. O vieno ar kito metodo pasirinkimas gali daug nulemti ir statistinės medžiagos pilnumą bei patikimumą, taip pat ir statistinio vaizdavimo būdus bei pasekmes.

7. Tiesioginis ir netiesioginis statistinis stebėjimas.

Apie tiesioginę statistiką, arba tiesioginį metodą statistikoj, kalbama tuo atveju, kai masės stovio arba reiškinių judėjimo stebėjimas daromas specialiai statistikos reikalui. Kai kuriuose statistikos vadovėliuose tokia statistika vadinama p i r m i n e arba p r i m a r i n e statistika (st. primo), kuri statoma, kaip antitezė, prieš a n t r i n ę arba s t. s e c u n d o, t. y. netiesioginę st. Kai kurie autoriai tiesioginės arba pirminės statistikos būtinu atributu dar laiko ir specialius statistikos organus. Tačiau ši pažiūra tenka laikyti ne visai teisinga. Yra teisybė, kad gyvenime, paprastai, tiesioginis statistikos stebėjimas daromas specialių statistikos organų betarpiškai, arba jų įsakymu. Tačiau ir ne specialūs statistikos organai vienu ar kitu atveju gali panorėti savam statistikos reikalui tiesioginį statistikos stebėjimą daryti. Ir tokių atsitikimų gyvenime netrūksta. Pav., mūsų geležinkelių valdyba daro sistemingą ir tiesioginį geležinkelių ridos bei transporto statistinį stebėjimą — pirmiausia savo orientacijai ir tik paskui tuos duomenis (jų dalį) patiekia Centriniam Statistikos Biurui, kuris juos skelbia kaip netiesioginės statistikos duomenis. Lietuvos Bankas savo orientacijai daro tiesioginį pasėlių stovio, derėjimų ir kai kurių prekių kainų bei pinigų rinkos statistinį stebėjimą. Kituose kraštuose tokius savam statistikos reikalui stebėjimus daro daugelis viešojo pobūdžio įstaigų ir net privatinės verslo įmonės, kurios su specialiais statistikos organais nieko bendra neturi.

a) Tiesioginė statistika. Jos esmė glūdi tame, kad, visų pirma, statistinis stebėjimas daromas betarpiškai stebint tyrinėjamąją masę ir, antra, šiuo stebėjimu gautoji skaitmeninė medžiaga betarpiškai tenaudojama tik bendrajai orientacijai, o ne kuriam administracijos aktui. Ta aplinkybė, kad vėliau, tokio statistinio stebėjimo duomenimis pasirėmus, būna ruošiamos ir daromos vienos ar kitos rūšies reformos, arba pavartojamos šiaip kurios nors rūšies bendrojo pobūdžio administracinės priemonės, tiesioginės statistikos esmės iškreipti negali, — jei visa tai taikoma ne statistikos stebėtiems vienetams paskirai, o visiems vienodai.

Pav., tiesioginio gyventojų surašymo duomenimis pasirėmus kai kuriose valstybėse nustatomi parlamento atstovų renkamieji koeficientai (kiek atstovų turi būti išrinkta į parlamentą). Kituose kraštuose toki koeficientai nustatomi mokesčių pajamoms arba išlaidoms, susietoms su atskirų tautybių mokymo arba atskirų tikybų kulto reikalais. Užsienių prekybos tiesioginio stebėjimo duomenimis gali būti pagrindžiamos vienos ar kitos rūšies muitų ar kt. prekybos bei ūkio politikos priemonės. Taip pat geležinkelių valdyba derina savo tarifų ir, aplamai, transporto politiką, atsižvelgdama statistinio stebėjimo bendrųjų duomenų, tačiau visai nesivaduoja gautaisiais duomenimis spręsdama individualiai kurios nors transporto siuntos ar kieno siuntinio konkretišką reikalą. Ir t. t.

b) Netiesioginė statistika. Jos esmė yra ta, kad, visų pirma, turimieji arba gaunamieji skaitmeniniai duomenys yra kilę ne iš specialinio stebėjimo, o prie kurios nors kitos progos ar kitam reikalui gautieji. Pav., visuose tardymų protokoluose pažymimas tardomųjų, o teismo bylose—bylininkų amžius, lytis, gyvenamoji vieta ir t. t. Mokesčių inspekcijoms patiekiamuose pareiškimuose būna surašomi duomenys apie įmonės ar asmens pajamas ir kt. Gyventojų registracijos įstaigose įrašomi duomenys apie atsikėlusių gyventi žmonių kilmę, amžių, verslą ir kt. Visi šios rūšies ir panašūs duomenys atitinkamų įstaigų arba organų tikrinami ir, jais pasirėmus, pirmiausia sprendžiamas individualiai konkretus reikalas, kuriam tie duomenys renkami buvo (pav., tarydmo protokolas siunčiamas teismo ar administracijos organams, teismo organai, spręs-

dami bylą, atsižvelgia bylininkų amžiaus, mokesčių inspekcijos nustato mokesčių dydį, ir t. t.). Ir tik vėliau, arba gretimai, turimieji duomenys vienu ar kitu metodu panaudojami statistikai.

c) Pranašumo bei pirmenybės klausimas vieną ar antrą statistinio stebėjimo metodą pasirenkant sprendžiamas atsižvelgiant statistinio stebėjimo uždavinių ir vietos aplinkybių. Aplamai, galima sakyti: jei netiesioginiu keliu gautoji statistikos medžiaga yra tiek pat tikra ir patikima, kaip ir būtų gauta tiesioginiu stebėjimu, ir jei tokios medžiagos statistikai panaudojimas gali sutaupyti lėšų ar laiko, nekenkiant statistinio vaizdavimo (medžiagos apdirbimo bei analizio) darbui, — tai tiesioginis statistikos stebėjimas nereikalingas. Tiesioginis statistikos stebėjimas gali būti nereikalingas arba netikslingas ir tais atvejais, kai toks stebėjimas vienu ar kitu atžvilgiu būtų pranašesnis (pav. duomenys greičiau gaunami ar kt.), tačiau pati masė arba reiškinys yra tokios rūšies, kad greitumo, statistinio pilnumo arba vaizdingumo reikšmė nėra labai didelė ir gali neapmokėti tiesioginio stebėjimo išlaidų.

Pirmajam atvejui gali būti pavyzdžiu mūsų miškų statistika. 1930 m. darant visuotinį žemės ūkio surašymą tiesioginiu metodu, buvo iš anksto nusistatyta šio metodo netaikinti valstybės laikomiems miškams, pasitenkinant netiesioginiais statistikos duomenimis iš miškų departamento. O priežastis ta, kad pilnesnių ir tikresnių duomenų, kokius galėjo turėti iš anksčiau miškų departamentas apie miškų kiekį, būklę ir sudėtį, — nebūtų buvę galima gauti ir tiesioginio stebėjimo keliu. Nes duomenų apie valstybinius miškus vistiek būtų buvę reikalinga klausti tų pačių valstybinės miškų administracijos organų; o reikalingus duomenis į surašomuosius lakštus tie organai būtų surašę taip pat tik tuos, kurie pas juos iš anksčiau buvo susikaupę. Taigi tiesioginis statistikos stebėjimas čia nieko nauja duoti negalėjo. Berods miškų administracija galėjo būti įpareigota iš naujo visus savo turimus duomenis patikrinti ar net specialiai statistikos reikalui miškus išmatuoti. Tačiau toks specialus statistikos reikalui miškų matavimas etc. buvo neįmanomas tuo trumpu laiku, kuris buvo nustatytas

žemės ūkio surašymui 1930 m. Taigi tiesioginio stebėjimo metodo taikymas miškų statistikai šiuo atveju tikresnių nė pilnesnių duomenų negalėjo duoti. Dėl to ir tiesioginio metodo pranašumas prieš netiesioginį šiuo atveju savaime atkrito.

Antrajam atvejui (kai tiesioginis st. stebėjimas pranašesnis, bet darbas neišsimoka) gali būti pavyzdžiu biudžetinių pajamų arba specialiai verslo pajamų statistinis stebėjimas. Nėra kalbos, kad tiesioginis biudžetinių pajamų stebėjimas statistikos reikalui duotų galimumo tas pajamas finansų statistikai daug tikslingiau ir gal greičiau sudoroti ir atvaizduoti, negu dabar netiesioginiu keliu gaunant duomenis per Finansų Departamentą. Taip pat verslo pajamų statistikai galima tikėtis gauti tikresnių ir pilnesnių duomenų, renkant juos betarpiškai statistikai, kai klausiamieji asmenys nesuinteresuoti tas pajamas mažinti arba slėpti, negu renkant tokius duomenis iš mokesčių inspekcijų, kur jie gaunami specialiai mokesčių apdėjimo reikalui. Tačiau statistikos praktika rodo, kad visais tais atvejais, kai klausiama apie individualinių pajamų dydį ar turtą, — net ir grynai statistikos reikalui, — visuomet atsiranda žymi dalis tokių žmonių, kurie yra linkę savo pajamas ar turto dalį mažinti, ir neatsiranda nė vienos bent kiek žymesnės grupės, kuri rodytų pajamų ar turto daugiau, negu faktinai turi. Dėl šios priežasties net ir idealiausias pajamų ar turto statistinis stebėjimas tiesioginiu metodu pilnai tikrų ir tikslių duomenų duoti negalėtų. Taip pat tiesioginio stebėjimo keliu nebūtų galima gauti pilnai patikimų žinių apie nusikaltimus arba nusikaltėlius: nes klausinėjant kiekvieną žmogų betarpiškai vis tiek nedaug atsirastų tokių, kurie norėtų prisipažinti nusikaltimą padarę net ir tuo atveju, kai apie tai dar niekas nežino. Dėl šios priežasties pajamų, nusikaltimų bei nusikaltėlių ir p. statistika būna dažniausiai paremta netiesioginiu stebėjimu (mokesčių inspekcijose, teismo bylose, policijos dokumentuose ir p.). Čia bent jau visi iš anksto žino, kad šios rūšies duomenys yra tikri tik tiek, kiek tiems organams pavyksta juos patikrinti.

Tiesioginio statistikos stebėjimo pranašumas savaime susidaro tais atvejais, kai stebimųjų arba reikalingų duomenų jokioj įstaigoj bet kuriam kitam reikalui surinktų nėra, arba esamieji duomenys nepilni ar nepatikimi, o tiesioginio stebėjimo

metodu galima tikėtis tikresnių duomenų gauti. Pav., duomenis apie javų derlius ar perteklius tik tiesioginio stebėjimo keliu galima gauti, nes šiaip jokia įstaiga šios rūšies duomenų nerenka. Duomenys apie žemės ūkių skaičių ir dydį galima rasti valsčių valdybų raštinėse, tačiau st. reikalui tie duomenys retai tenaudojami dėl šių priežasčių: a) ūkių sąrašai būna sudaryti grynai juridiniais pažymiais, neatsižvelgiant faktinojo valdymo ir naudojimo aplinkybių, b) tuose sąrašuose esamieji duomenys surašyti juridinio valdymo pakeitimo metu, ir, paprastai, mažai atsižvelgiama vėliau įvykusių pakitimų, c) visos eilės statistikai svarbių duomenų (apie inventorių, trobesius, darbo jėgas ir kt.) tuose sąrašuose paprastai nebūna, arba būna sužymėta ne vienu metu, ir d) iš valsčių valdybose esančių ūkių sąrašų statistikai duomenų išrinkimas ir jų surašymas nedaug mažiau lėšų ir laiko atimtų, kaip tiesioginis stebėjimas ir žymėjimas vien statistikos reikalui.

Be to, būna atsitikimų, kai bet kurios rūšies įstaigose esanti statistinė medžiaga yra mažai patikima, nežiūrint į jos tariamą pilnumą bei nuolatinę kontrolę. Pav., jau senai įvairiuose kraštuose pastebėta, kad darant gyventojų statistinį stebėjimą tiesioginiu ir netiesioginiu metodu tais pačiais metais, duomenys gaunami žymiai skirtingi. Ir tas skirtumas pasikartoja gana dėsningai: tiesioginio stebėjimo keliu gautasis gyventojų skaičius išeina visuomet mažesnis už netiesioginės statistikos duomenis, paremtus prisirašymo arba prisimeldavimo aktais. Net ir tų skirtumų proporcija dažniausiai pasikartoja vidutiniškai kaip 4:5, t. y., tiesioginio stebėjimo keliu gautasis gyventojų skaičius būna maždaug apie 20% mažesnis už skaičių, gautą netiesioginio stebėjimo keliu. Ir Lietuvoj 1920 m. netiesioginio stebėjimo keliu nustatant bendrąjį gyventojų skaičių ryšyj su numatomais rinkimais į Steigiamąjį Seimą buvo rasta gyventojų daug daugiau, kaip 1923 m. tiesioginio stebėjimo keliu darant surašymą. Priežastis čia, kaip ir kituose kraštuose, būna dažniausiai ta, kad persikeliant iš vietos į vietą prirašymo arba vad. primeldavimo (įregistravimo) žiūrima stropiau, kaip išrašymo (išregistravimo); taip pat ir numirusieji ypač valsčiuose, paprastai, dar gana ilgai būna gyvųjų sąrašuose.

Taigi tiesioginio ar netiesioginio stebėjimo metodo pranašumas pareina nuo stebimos masės pobūdžio ir nuo įvairių vietos bei laiko aplinkybių, taip pat nuo paties stebėjimo tikslo bei pasistatytų uždavinių.

d) Tiesioginės ir netiesioginės statistikos metodų kombinacija arba junginys atskirais atvejais taip pat gali būti naudojama. Pavyzdžiui, vad. civilinių aktų arba metrikacijos įstaigose visai tiksliai registruojami gimimai, mirimai ir vedybos lyties, amžiaus, tikybos ir kilmės pažymiais. Taip pat muitinėse registruojamos per sieną einančios prekės. Visai savarankiškas ir betarpiškas specialių statistikos organų stebėjimas šiais atvejais nei pilnesnių, nei tikresnių duomenų patiekti negalėtų. Tačiau netiesioginis stebėjimas geriausiu atveju duotų žymiai pavėlintus duomenis, o dažnai ne tais pačiais pažymiais sužymėtus. Tokiomis aplinkybėmis gali būti tikslinga tiesioginį ir netiesioginį stebėjimą sujungti ta prasme, kad, pav., civilinio stovio akto surašytojas įpareigojamas tame akte pažymėtus duomenis tam tikra griežtai nustatyta tvarka ir statistikos reikalui į specialinį lakštą įrašyti bei nustatytais terminais statistikos organams patiekti. Tas pats gali būti daroma ir muitinėse su užsienių prekybos prekėmis, mokesčių inspekcijose su patiekiamais duomenimis apie verslo įmonių būklę, mokesčių mokėtojų pajamas ir t. t.

Dažnai, tačiau, šios rūšies statistinio stebėjimo organizacija eina ligi to, kad išeina jau ir savarankiškas tiesioginis statistikos stebėjimas. Pav., pas mus dabartiniu metu natūrinio gyventojų judėjimo statistika yra suorganizuota taip, kad civilinės padėties aktų sustatytojai (klebonai ir kt.) yra įpareigoti specialiai statistikos reikalui surašyti surašomuosius lakštus apie kiekvieną gimimą, mirimą ir t. t. tam tikrais pažymiais visai nepriklausomai nuo duomenų, reikalingų civilinio akto formalumams atlikti (pav., klausiama apie verslą, raštingumą ir kt.). Be to, pati statistikos surašomoji kortelė turi būti surašoma anksčiau, negu vykdomas su metrikos sudarymu siejamas aktas (jungtuvės, krikštas ar kt.). Ir muitinių valdininkai apie kiekvieną prekę statistikos reikalui duomenis privalo surinkti nepriklausomai nuo to, ar prekė muitinama ar ne muitinama, ir ar muitinimo reikalui tokios žinios reikalingos ar ne-

reikalingos. Tuo būdu patsai statistinis stebėjimas išeina jau tiesioginis, tik atliekamas ne specialių statistikos organų, o kitų, visai kitam reikalui sudarytų įstaigų, kurios, tačiau, statistinį stebėjimą daryti yra įpareigotos (ir atlieka) lygiai taip, kaip atliktų ir betarpiški statistinio stebėjimo organai.

8. Individualinis ir kolektyvinis statistikos stebėjimas.

a) Individualinis statistikos stebėjimas vadinamas toks st. surašymas, kur kiekvienas skaitomasis vienetas į surašymo lakštą (kortelę) atskirai įrašomas, čia pat pažymint visas surašomąsias žymes bei pažymius. Pati surašymo technika gali būti nevienoda ir vykdoma arba kortelių sistema arba sąrašų sistema. Kai kiekvienam skaitomajam vienetui surašoma atskira kortelė, tai toks surašymas priklauso kortelių sistemos, o kai į vieną surašomąjį lakštą įrašoma keletas skaitomųjų vienetų iš eilės, tai vadinama sąrašų sistema.

Lietuvoj pirmasis visuotinis gyventojų surašymas 1923 m. ir pirmasis visuotinis žemės ūkių surašymas 1930 m. buvo padaryti kortelių sistema. 1897 m. buvęs gyventojų surašymas Rusijoj ir Lietuvoje, taip pat 1925 m. ir vėlesnieji gyventojų surašymai Vokietijoje buvo daromi sąrašų sistema. Čia į vieną surašomąjį lakštą iš eilės buvo įrašoma tiek skaitomųjų vienetų, kiek skaitomąją dieną buvo rasta viename surašomajame vienete (šeimoje arba viename ūkyje) asmenų.

Kuri iš šių dviejų sistemų yra pranašesnė — pareina beveik išimtinai nuo to, ar statistinio stebėjimo (surašymo) duotoji medžiaga yra apdirbama (klasifikuojama) rankomis ar naudojamos tam tikros, statistinę medžiagą klasifikuojamos, mašinos bei aparatai. Aplamai, kortelių sistema tuo yra patogi, kad, medžiagą apdirbant, kortelės, kurios šioj sistemoj atstovauja skaitomuosius vienetus, gali būti įvairiausiais pažymiais saistomos bei grupuojamos, ir kiekvienos grupės bendrasis skaičius surandamas korteles rankomis skaičiuojant. Sąrašų gi sistemoj kiekvieno skaitomojo vieneto pažymiai iš sąrašo turi būti paimti ir vienu ar kitu būdu atžymėti kuriame nors kitame vieno pažymio sąraše ir tik paskui jau gali būti sumuojama. Statistinės medžiagos apdirbimo praktikoj yra patirta ir nusta-

tyta, kad statistinę medžiagą rankomis apdirbant iš sąrašų, geriau išsimoka ir laiko susitaupo, jei iš bendrojo sąrašo skaitomieji vienetai su visais pažymiais surašomi į korteles kiekvienas atskirai lygiai taip pat, kaip skaitomieji vienetai į korteles surašomi kortelių sistema. Vadinasi, statistinę medžiagą rankomis apdirbant — statistinis stebėjimas kortelių sistema yra parankesnis. Sąrašų sistema tik tuo atveju parankumo atžvilgiu gali būti lyginama su kortelių sistema, kai sąraše skaitomieji vienetai įrašomi taip, kad vėliau (medžiagą apdirbant) tie vienetai duodasi lengvai vienas nuo antro atskiriami (atkerpami arba atplėšiami), t. y., praktiškai naudojami taip pat, kaip ir kortelių sistemos stebėjimu gautoji medžiaga. Bet jei sąrašai čia paminėtu būdu ir forma sudaryti, tai savo esmėje jie nuo kortelių sistemos beveik ir nesiskiria.

Tačiau kortelių sistema visiškai pasidaro nereikalinga, kai statistinė medžiaga apdirbama (klasifikuojama) mechaniškai, t. y., naudojant specialines vienos ar kitos sistemos mašinas. Nes mašinomis statistikos medžiagą apdirbant negali būti betarpiškai naudojamos tos pačios kortelės, kurios yra naudojamos st. medžiagai renkant (surašinėjant). Šios kortelės, visų pirma, yra labai nedidelio formato (paprastai apie 185×80 mm). Be to, jos negali būti nei sulankstytos, nei suglamžytos (paprastai būna kartoninio popierio arba net celiuliozinės). Į šias korteles surašomųjų žymių kiekvienas pažymys turi milimetro griežtumu apibrėžtą savo vietą, kuri kortelėje atžymima skylutės pradūrimu. Taip subadyta kortelė (su tiek skylučių, kiek iš viso nusistatoma apdirbti skaitomųjų pažymių) įdedama į korteles skirstomosios mašinos tam tikrą stalčių, iš kurio jau automatiškai kortelės, kiekviena sulig savo turimų skylučių paskirstymu, eina tam tikrais takais atskiromis grupėmis. Kiekvienos tokios grupės vienetai, taip pat prie visų grupių bendrieji vienetų skaičiai, taip pat automatiškai suskaitomi ir bendrosios sumos parodomos. Prie šitokios statistinės medžiagos apdirbimo sistemos, statistinės medžiagos rinkimas kortelių sistema savo pranašumo nustoja (žiūr. p p. 90—94).

b) Kolektyvinis st. stebėjimas bus tuomet, kai į vieną surašomąjį lakštą iš karto įrašomas visas skaitomųjų vienetų kolektyvas, kuris čia pat suskirstomas surašomųjų žymių pažymiais. Toks surašomasis vienetas čia gali

būti, jei tai gyventojų surašymas, visa šeima ar visas sodžius, ar kuris kitas žmonių kolektyvas. Šis kolektyvas jau iš karto suskirstomas į tiek grupių, kiek yra nusistatyta žymėt pažymių (lyties, amžiaus, tautybės, tikybos, verslo ar kt.).

c) Pranašumo arba pirmenybės klausimas. Kuris statistinio stebėjimo metodas (individualinis ar kolektyvinis) yra tinkamesnis, pareina nuo įvairių aplinkybių ir statistinio stebėjimo uždavinių. Statistika privalo parodyti masės (arba masinių reiškinių) charakteringas ypatybes. Bet kokias ypatybes statistikos metodais atvaizduoti — galima tik skaidant skaitmeninę medžiagą. Vadinasi, kiekvieno st. darbo uždavinys — st. grupių sudarymas. O jeigu taip, tai ir statistinis stebėjimas, grynai teoretiškai imant, turėtų būti galimas tiek individualiniu, tiek kolektyviniu metodu. Tik reikalinga, kad, kolektyviniu metodu darant statistinį stebėjimą, jau iš karto būtų stebimi tik toki kolektyvai (tokios grupės), kurie, bendrai paimti, nepaslėptų, arba neiškreiptų, charakteringųjų masės (masinio reiškinio) ypatybių.

Pavyzdžiui galime paimti gyventojų statistinį stebėjimą. Jei mes iš anksto nusistatysime, kad prie jokių aplinkybių statistinis vaizdavimas nebus daromas smulkesniais kaip sodžiaus ploto pažymiais, ir jeigu, be to, iš anksto apibrėšime skaitomųjų žymių ir pažymių skaičių ir turinį (lyties, amžiaus, tautybės ir kt. pažymius), tai statistinio stebėjimo duomenų vaizdavimo bendrieji rezultatai galės būti tie patys, — nepriklausomai nuo to, ar kiekvienas skaitomasis vienetas (gyventojas) bus stebimas ir žymimas atskirai (griežtai individualiniu metodu), ar bus tik individualinis stebėjimas, o kolektyvinis (visam sodžiui ar kt. kolektyvui bendrai) žymėjimas.

Yra tokios rūšies masių arba masinių reiškinių, kur individualinis stebėjimo žymėjimas beveik niekuomet ir nepraktikuojamas. Pav., gyvulių statistinį stebėjimą darant niekuomet nereikalaujama, kad gyvulių laikytojai kiekvienam gyvuliui (arkliui, galvijui, kiaulei ir t. t.) surašytų atskirai surašomąjį lakštą, arba atskirai aprašytų kiekvieno gyvulio metus, veislę bei rūšį ir kitus pažymius. Paprastai iš gyvulių laikytojų reikalaujama pasakyti arba pažymėti, kiek iš viso jie turi arklių, galvijų, avių, kiaulių, vištų ir kt., kiek yra arklių ar galvijų

veislės ligi tam tikro amžiaus, kiek senesnių ir t. t. Taigi beveik visuomet žinios teikiamos jau apie visą kolektyvą, arba atskiras jo grupes, kartu. O jeigu atsitiktinai pasitaiko, kad kas nors turi tik vieną kokį gyvulį ir dėl to surašomajame lakšte įrašomi duomenys tik apie vieną skaitomąjį vienetą, tai čia tik atsitiktinis nuo normos nukrypimas, kuris neiškreipia kolektyvinio stebėjimo esmės.

Aplamai, tačiau, imant pirmenybę tenka atiduoti individualiniam statistinio stebėjimo metodui, — išskiriant tuos atvejus, kai individualinis statistinio stebėjimo metodas dėl vienų ar kitų priežasčių būtų neįmanomas arba visai nereikalingas.

Dalykas toks, kad kolektyvinis statistinio stebėjimo metodas yra būtinai reikalingas griežto aprėžimo turinio ir skaičiaus atžvilgu visų skaitomųjų žymių ir pažymių dar prieš stebėjimo pradžią. O toksai aprioristinis aprėžimas ne tik sunkiai įmanomas visais tais atvejais, kai stebimoji masė nelabai gerai pažįstama (patsai gi statistinis stebėjimas daromas, kad būtų galima pažinti), bet ir pavojingas; nes ne patyrimu paremtas ypatybių aprėžimas gali iškreipti stebimos masės charakterį.

Pavyzdžiui imkime pas mus buvusius gyventojų ir žemės ūkių surašymus. Visos žymės ir pažymiai, kuriais gyventojų arba ūkių surašymo medžiaga vėliau buvo apdirbama, galima buvo aprėžti ir nustatyti dar prieš surašymą. Bet surašomieji lakštai kolektyviniam stebėjimo metodui jau iš anksto būtų privalėję turėti atskiras skiltis kiekvienai gyventojų amžiaus grupei, kiekvienam gyventojų tautybės, tikybos ir verslo pavadinimui ir t. t. O atskirai pažymėtų pavadinimų gyventojų surašymo knygose (Lietuvos Gyventojai) mes randame: tautybėms — 25, tikyboms — 30, verslams — 63, ir t. t. Žemės ūkių surašymas kolektyviniu metodu būtų buvęs reikalingas surašomųjų lakštų su keliais šimtais atskirų skilčių. Jei ir galima būtų iš anksto visas reikalingas skiltis ir pavadinimus numatyti, vis dėlto reikėtų abejoti, kad tie patys surašinėtojai būtų sugebėję tokioj daugybėj skilčių orientuotis ir tiksliai kiekvieną skaitomąjį vienetą į atitinkamą skiltį įrašyti. Aplamai, yra jau senai žinoma taisyklė: juo daugiau yra lakšte atskirų

skilčių ar pavadinimų ir juo dažniau atskiri pažymiai žymimi ne žodžiais, o skaičiais ar kuriais ženklais, — tuo daugiau įsiskverbia stebėjimo paklaidų.

Kolektyvinio stebėjimo metodui pirmenybė prieš individualinį metodą galima duoti tik tais atvejais, kai dėl vienų ar kitų priežasčių statistinis stebėjimas daromas specialiai iš anksto parinktiems arba aprėžtiems pažymiams nustatyti ar kai pirmoj eilėj statomas ne visapusiškas statistikos stebėjimas, o dalinis arba orientacinis, skubiam reikalui esant, surašymas. Pavyzdžiu gali būti reikalas skubos keliu pavasarį sužinoti, kiek pas ūkininkus yra jauniklių gyvulių, arba kiek, maždaug, atskiroms kultūroms yra pasėlių. Individualinis visuotinis stebėjimo metodas šiais atvejais ir labai brangiai kaštuotų, ir perdaug ilgai užsitęstų. Kolektyviniu gi metodu renkant duomenis ištisiems sodžiams per valsčių valdybas, darbas žymiai lengvėja ir pasigreitina. O duomenų tikrumas ir individualinio stebėjimo metodu šiais atvejais visai nelaiduojamas. Šiais sumetimais, pav., kai kuriuose kraštuose duomenys apie pasėlių plotus kolektyviniu metodu ir būna renkami. Tačiau ir tai dar galima abejoti, ar, užuot ištisinio arba visuotinio stebėjimo visų kolektyvų, negalėtų būti tikslingiau ir sėkmingiau pavartotas individualinis stebėjimas reprezentaciniu metodu.

9. Ištisinis ir dalinis st. stebėjimas bei reprezentacinis metodas.

a) Ištisinis st. stebėjimas apima visus be išimties stebimos masės vienetus, kurie ir surašomi be jokių praleidimų. Pav., jei statistikos stebimi gyventojai, tai į surašomuosius lakštus įrašomi visi be jokios išimties gyventojai, jei gyvuliai — tai visi gyvuliai, jei trobėsiai ar ūkiai — tai visi trobėsiai, ūkiai ir t. t. Vadinasi, ištisinis st. stebėjimas privalo visuotinio surašymo.

b) Dalinis st. stebėjimas apima ne visą masę, o tik kurią nors masės dalį. Tačiau patsai stebėjimas daromas tuo tikslu, kad iš surašytosios dalies apie visą masę būtų galima spręsti. Pav., darant dalinį gyventojų surašymą, gyventojai gali būti surašinėjami ne visame plote, o tik vienoj

antroj apskrityje, arba kiekvienos apskrities tik viename valsčiuje, ar kt. Tuo būdu surašinėjimui parinktieji rajonai (apskritys, valsčiai ar kt.) reprezentuotų visumą — visą valstybę. Bet surašinėjama gali būti taip pat tik masės dalis ne paskirai paimtuose rajonuose, o visame plote vienu ar kitu metodu paimti skaitomųjų vienetų reprezentantai. Tuo būdu gautieji duomenys laikomi tipiškais visai stebimai masei ir, paprastai, būna apibendrinami. Pav., derėjimų statistikai duomenys paprastai renkami ne visuotinio surašymo keliu (apklausiant visus be išimties ūkininkus ir sužymint kiekvieno iš jų gautąjį derlių į surašomuosius lakštus), o pasirėmus tik vieno — antro ūkininko patiektais duomenimis iš visos apylinkės (rajono, seniūnijos, valsčiaus). Taip pat duomenys prekių kainų statistikai gaunami ne ištisiniu stebėjimu visų be išimties parduoties atsitikimų, o stebima ir surašoma tik kurios nors savaitės, ar net vienos dienos per mėnesį, dažniausia mokamos kainos. Ir t. t.

c) Reprezentacinis metodas statistikoj yra pažįstamas įvairiuose variantuose. Vienas iš tokių variantų — tai jau minėtas dalinis arba nepilnas st. surašinėjimas. Dalinio st. surašinėjimo esmė yra ta, kad st. masė suskirstoma dalimis ir šitoj dalyj (rajone, jei tai stovio st., skaitomojo periodo aprėžtame laikotarpyje — jei judėjimo st.) visi skaitomieji vienetai yra stebimi ir į surašomuosius lakštus įrašomi lygiai taip pat, kaip ir per ištisinį stebėjimą (visuotinį surašymą).

Statistinės anketos — tai kitas reprezentacinio metodo variantas. Anketinis st. stebėjimas savo išviršine forma gali beveik nieku nesiskirti nuo ištisinio st. stebėjimo: nes anketos, t. y. surašomieji lakštai, gali būti išsiuntinėti visiems be išimties st. masės vienetus atstovaujantiems subjektams, kurie reikalingus duomenis į surašomuosius lakštus (anketas) įrašo. Skirtumas, tačiau, tas, kad anketas užpildo tik savanoriai be jokios prievartos ir be kontrolės. Be to, anketinio stebėjimo praktikoj, paprastai, leidžiama arba net pageidaujama įrašyti savo pastabas, paaiškinimus ir net pageidavimus. Savaime, t. y. teorijoj, yra galima, kad anketinis st. stebėjimas nieku nesiskirtų nuo ištisinio stebėjimo arba visuotiniojo surašymo: reikalinga tiktai, kad anketos lakštai būtų surašyti visiems masės vienetams tokiu pat tikslumu bei rūpestingumu, kaip per visuotinį surašymą. Tačiau ir šiuo atveju duotieji statistiniai duomenys turėtų būti laikomi mažiau patikimais, nes, nesant kiekviename statistikos stebėjime būtinos kontrolės, niekas negalėtų pasakyti, ar tikrai visi skaitomieji

vienetai yra surašyti; taip pat nebūtų galima žinoti, ar visos surašomosios žymės su pažymiais teisingai buvo suprastos ir tiksliai atžymėtos. Aplamai, tiek dalinis st. surašymas, tiek anketinis st. stebėjimas statistinio metodo reikalavimus negali pilnai patenkinti dėl to, kad nesiduoda griežtai kontroliuojamas, ir dėl to beveik niekuomet negalima apibrėžti galimos statistinio stebėjimo paklaidos ribas. Kad bet koks statistinis stebėjimas būtų patikimas, yra reikalinga, kad jis pasiduotų kontrolei. Tuo atžvilgiu patikimesnius duomenis garantuoti gali atsitiktinių ar tipiškųjų reprezentantų statistinis stebėjimas.

d) Atsitiktinių reprezentantų statistinis stebėjimas daromas tuo būdu, kad iš anksto nusistatoma tik „kelintąjį" stebimos masės skaitomąjį vienetą stebėti ir statistikos reikalui žymėti. Pav., galima nusistatyti, kad namų ar butų statistikai aplankoma tik kiekvieną penktąjį, dešimtąjį ar dvidešimtąjį namą ir jo charakteringas ypatybes įrašoma į surašomąjį lakštą. Taip pat, užuot surašinėti ištisai visus žemės ūkius arba jų derlius, galima žymėti penktojo, dešimtojo, penkiasdešimtojo ar kt. ūkio derlius. Ir jei tik bendrasis namų ar ūkių skaičius pakankamai didelis, tai, einant didžiųjų skaičių dėsniu, galima tikėtis, kad kiekvienas penktasis, dešimtasis ar dvidešimtasis ūkis, aplamai imant, atitiks ir likusiųjų namų ar ūkių ypatybes, ir kad galutinieji duomenys, išskaičiuoti proporcionaliai, nuo tikrenybės išeis perdaug nenutolę.

Tačiau toksai reprezentantų parinkimo metodas gali būti tikslingas tik tais atvejais, kai likimo skirtieji stebėjimo vienetai yra vien stebėjimo objektai, kurių ypatybės kiekvienam aiškiai matomos, ir pašalinio stebėtojo, t. y. surašinėtojo, gali būti pastebimos. Tačiau, jei stebimos masės skaitomieji vienetai turi ypatybių, kurios pašaliniam stebėtojui nematomos, tai atsitiktinių reprezentantų parinkimas komplikuojasi. Pav., pasėlių stovio arba derėjimų statistikai stebimieji ūkiai gali būti imami ir atsitiktinai (dešimtasis, šimtasis arba kiekvieno sodžiaus pirmasis, antrasis ar kt. ūkis). Tačiau pasėlių stovis yra reikalingas nuolatinio stebėjimo vietoje; taip pat ir derlius tegali būti tiksliau stebimas ir atžymimas tik paties ūkininko. O ne kiekvienas ūkininkas yra vienodai pastabus, be to, net ne kiekvienas ūkininkas gali būti pakankamai raštingas, kad savo stebėjimo duomenis galėtų sistematingai sužymėti ar atlikti visus kitus derlių statistikos stebėtojui privalomus darbus.

Taigi čia jau reikalinga ir stebimų ūkių valdytojų, arba stebėtojų, kvalifikacijos, kurias nustatyti gali tik atsargus parinkimas, o ne aklas likimas. Dėl šios priežasties ir derėjimų statistikai reprezentaciniu metodu patys stebėtojai (pas mus vadinami žemės ūkio statistikos korespondentai) parenkami iš tokių vietos ūkininkų, kurių ūkiai maždaug yra tipiški vietos sąlygoms, o patys ūkininkai pakankamai inteligentiški technikiniam stebėjimo ir atžymėjimo darbui atlikti. Taigi, gana dažnai parenkami tipiškieji reprezentantai atrankos keliu.

e) Tipiškųjų reprezentantų parinkimas jau iš karto turi būti daromas atsižvelgiant vad. tipiškiems dydžiams statomų teorijos reikalavimų. Dalykas toks, kad atsitiktiniai masės reprezentantai — jei tik pakankamai didelis tokių reprezentantų skaičius — visais atžvilgiais gali ir privalo būti traktuojami lygiai taip, kaip ir kiekvienas atskiras skaitomasis vienetas ištisiniame stebėjime. Pav., jeigu būtų nusistatyta kiekvieną dešimtąjį ūkį laikyti visos ūkių masės reprezentantu, ir jeigu, tuo būdu stebint, būtų surašyta iš viso 1000 ūkių, ir tame skaičiuje atsirastų 10 ūkių po 5 ha, 20 ūkių po 10 ha, 50 ūkių po 20 ha, 10 ūkių po 30 ha, 9 ūkiai po 50 ha ir 1 ūkis po 100 ha, tai nebūtų jokio pagrindo šią ūkių dydžio proporciją keisti bendrąjį ūkių skaičių nustatant. Taip pat ir tuose ūkiuose pažymėtus pasėlių ploto, inventoriaus ar derėjimų duomenis reikėtų taikyti griežtai atsitiktinių reprezentantų duomenimis nustatytus.

Kitaip jau reikėtų elgtis parenkant vad. tipiškuosius reprezentantus. Pav., šiuo atveju iš viso būtų galima abejoti, ar 100 ha ūkiai, jei jie bendroje ūkių masėje tik 1% tesudaro, dar būtų reikalingi savojo reprezentanto statistiniam stebėjimui: nes bendrieji duomenys apie derlių, inventorių ar kt. vistiek negalėtų žymiai pasikeisti, jei tokio nedidelio ūkių skaičiaus visai nebebūtų atsižvelgiama.

Apskritai imant, tipiškųjų reprezentantų statistinis stebėjimas turi būti bazuojamas „tipais“, t. y. masės daugumą atstovaujančiais reprezentantais — atsižvelgiant vidurkių ir statistinių dydžių arba eilučių dispersijos. Iš to seka, kad tipiškųjų reprezentantų statistinio stebėjimo duomenys gali, o

tam tikrais atvejais net ir privalo, būti koriguojami, jei yra pagrindo tarti, kad dėl vienų ar kitų priežasčių į tuos duomenis įsiskverbia ir anormalinių, t. y. netipiškų, veiksnių įtaka.

f) Pranašumo bei pirmenybės klausimas vieną ar antrą metodą pasirenkant šiandien dar ne visų statistikos atstovų vienodai sprendžiamas. Aplamai, galima sakyti, kad beveik visi griežtieji vad. istoriškai aprašomosios statistikos šalininkai pirmenybę atiduoda ištisiniam stebėjimo metodui arba visuotiniam surašymui. Reprezentacinį metodą jie teprileidžia bei pateisina tik tais atvejais, kai dėl vienų ar kitų priežasčių ištisinis stebėjimas, arba visuotinis surašymas, nėra galimas.

Grynai teoretiškai imant tenka sutikti, kad duomenys, gautieji visuotinio stebėjimo keliu, kai kiekvienas skaitomasis vienetas įrašomas į surašomąjį lakštą, galėtų būti tikresni ir teisingesni už duomenis, kurie skaičiuoti pasirėmus tik daliniu masės stebėjimu. Prieš šitą skaičiavimo teorijos argumentą nieko nebūtų galima pastatyti ir už reprezentacinį metodą statistikoj nieko nebūtų galima pasakyti, jei visais atvejais ištisinio statistikos stebėjimo metodas būtų galimas praktikoj įgyvendinti ir, be to, jeigu visais atvejais visuotinis surašymas laiduotų visai tikrus ir pilnai patikimus duomenis. Tačiau gyvenimo praktikoj nevisuomet taip būna. Pav., beveik visi ūkininkai apie savo pasėlių plotus sprendžia remdamiesi ne tiksliais jų plotų matavimo duomenimis, o „iš akies“ — arba sulig tuo, kiek sėklos pasėta. Taip pat ir apie gautąjį derlių ūkininkas sprendžia nevisuomet iš pilnų iškulto javo svorio ar saiko duomenų. Taigi šiais atvejais negali būti laiduota, kad ištisinis stebėjimas arba visuotinis surašymas galėtų griežtai tikslius ir tikrus duomenis duoti. Ir visus ūkininkus apklausinėjant apie pasėlius ir derėjimus, vis dėlto duomenys geriausiu atveju bus tik apytikriai, t. y. spėtiniai: nes geriausiu atveju viena ūkininkų dalis tik kurią nors savo pasėlių plotų ar derlių dalį bus tiksliai išmatavusi, o, aplamai, vadovausis „iš akies ir atminties“ prisiminimais arba iš analogijos, t. y. tik apytikriai spėjamais skaičiais.

Be to, gali būti atsitikimų, kai visuotinio stebėjimo keliu gautieji duomenys bus netikri ir nepilni dėl to, kad klausiamųjų

asmenų dauguma dėl vienų ar kitų priežasčių gali nepanorėti visai tikrus duomenis patiekti (pav., apie pajamas ar net ir derlius daug kas nenorės pilnų ir tikrų duomenų patiekti dėl to, kad bijos didesnių mokesčių ar kt.); tokiais atvejais kai kada iš tiksliai ir atsargiai parinktų reprezentantų galima teisingesnius duomenis gauti.

Taip pat gali būti atsitikimų, kad ištisinis stebėjimas techniškai arba fiziškai neįmanomas. Pav., ištisinio stebėjimo keliu galima nustatyti ir surašyti miškų plotą ištisai visus miškus matuojant, nors toks darbas yra labai brangus ir ilgas; tačiau miškų tūrį ištisinio stebėjimo keliu nustatyti neužtektų pajėgų ir visus šalies gyventojus prie to darbo patraukus. Tuo tarpu reprezentaciniu metodu miškų tūris su gana didele tikslumo tikimybe surandamas ir neperdidelio skaičiuotojų kiekio per, palyginti, neilgą laiką.

Iš kitos, tačiau, pusės yra masių arba masinio pobūdžio reiškinių, kur be ištisinio stebėjimo ir be visuotinio surašymo bent kiek patikimesnių duomenų gauti negalima. Pav., apie vad. „sunkiosios“ pramonės įmonių skaičių ar apie jų gamybos rūšį (taip pat apie kai kurias kitas ypatybes) tik ištisinio stebėjimo keliu galima tikėtis bent kiek patikimesnių duomenų gauti. Nes čia skaitomieji vienetai savo žymėmis yra tiek skirtingi, o vienodomis žymėmis įmonių tiek gali būti nedaug (jei iš viso jų būna), kad bet koks dalinis surašinėjimas nėra galimas dėl statistinės masės, taigi ir ją atitinkančių tipiškų reprezentantų, stokos. Taip pat sunku tikėtis net apytikrių duomenų gauti reprezentaciniu metodu apie bankus ir jų operacijas: nes bankų iš viso yra nedaug, ir kiekvienas iš jų daugiau turi savaimingų ypatybių, ne kaip bendrojo pobūdžio pažymių. Šiuo atžvilgiu didesnio vienodumo tegalima rasti tik vad. liaudies bankeliuose arba smulkiojo kredito draugijose.

Aplamai, galioja dėsnis: juo mažesnė yra stebimoji masė ir juo ji nevienodesnė, tuo mažiau gali būti pritaikomas reprezentacinis metodas; ir atvirkščiai: juo stebimoji masė didesnė ir juo ji vienodesnė, tuo sėkmingiau gali būti pritaikomas reprezentacinis metodas.

g) Elementarinis reprezentacinio metodo pagrindimas.

Matematinis reprezentacinio metodo pagrindimas, taip pat reprezentaciniu metodu gautųjų duomenų patikimumo ribos aprėžimas bei patikrinimas, yra reikalingas tikimybinio skaičiavimo tiekiamųjų priemonių bei metodų. Tačiau patys tokio tikimybinio aprėžimo principai duodasi ir paprastais samprotavimais suvokiami.

Sutarkime, kad st. masės skaitomųjų vienetų x_i bendrasis skaičius mums nėra žinomas, tačiau yra žinoma, kad šitoje masėje viena dalis vienetų turi pažymį A, o kita B (tarsime — vyrai ir moterys, arba raštingi ir neraštingi). Jei, be to, sutarsime reprezentantais imti kiekvieną dešimtąjį, o visus be išimties vienetus griežtai skaičiuosime po vieną iš eilės kaip x_1, x_2, x_3, ..., x_n, tai bendrąjį masės vienetų skaičių galėsime nustatyti visai tiksliai: nes iš viso skaitomųjų vienetų bus tiek, kiek sudarys pilnų dešimtinių skaičius, padaugintas reprezentantų skaičiumi plius liekana iš paskutinės nepilnos dešimtinės. Jeigu, be to, sutarsime x-inės eilutės reprezentantus žymėti $\dot{x}_i$, tai, imant kiekvieną dešimtąjį, pirmasis reprezentantas $\dot{x}_1=x_{10}$, antrasis $\dot{x}_2=x_{20}$, $\dot{x}_3=x_{30}$ ir t. t. ligi paskutiniojo, o nepilnos dešimtinės liekaną žymėsime z. Tuo būdu bendroji x_i vienetų suma visoj masėj bus lygi:

$$x_1+x_2+x_3+....+x_n=\Sigma x_i=$$
$$=10\dot{x}_1+10\dot{x}_2+...+10\dot{x}_m+z=10\Sigma\dot{x}_i+z.$$

Jei turėsime dvi mases A ir B, iš kurių pirmoji A=300, o B=305, tai reprezentacinis metodas iš dešimtinių reprezentantų pagal formulą $X=10\ \Sigma\dot{x}_i+z$ duos:

$$A=10a_1+10a_2+10a_3+z=10(10+10+10)+0=300.$$

Tuo pačiu principu:

$$B=10\ \Sigma b_i+z=10(10+10+10)+z=300+z.$$

Bet šiuo atveju z=5, taigi ir formula duos:

$$B=10\,(\Sigma b_i)+z=300+5=305.$$

Vadinasi, atsitiktinių reprezentantų stebėjimas įgalina tam tikrose sąlygose iš anksto nežinomos X masės vienetų skaičių tokiu pat griežtumu nustatyti, kaip ir ištisinis stebėjimas (visuotinis surašymas).

Čia pat pravartu susipažinti su liekanos reikšme. Jeigu reprezentantams imamos masės koeficientui yra nustatytas pastovus (nekintamais intervalais pasikartojantis) eilutės narys (dešimtasis, šimtasis, ar kt.), tai z=0 bus visais tais atvejais, kai X:k duos sveikąjį skaičių (be trupmenos). Vadinasi, jeigu atsitiktinių reprezentantų koeficientas k=10, o N=30, 100, 1000 ar kitas kuris skaičius su 0 gale, tai iš reprezentantų skaičiaus N masės vienetų skaičiui surasti galiojanti formula

$$N=k\,\Sigma\dot{x}_i+z \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (I)$$

redukuojama į

$$N=k\,\Sigma\dot{x}_i \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (Ia)$$

Nes: jeigu k=10, o N=30, 100 ar kt.,

tai visais atvejais

$$\Sigma x_i = k\Sigma\dot{x}_i$$

Be to, z reiškinio, t. y. liekanos, reikšmė reprezentuojamai masei mažėja stebimos masės vienetų skaičiui didėjant ir, be to, reprezentacijos koeficientui k mažėjant. O tai yra svarbi aplinkybė. Pav., jeigu k=10, tai stambiausia liekana z gali būti 9. Bet jeigu paimsime k=5, tai z negalės būti didesnė kaip 4. Iš to seka, kad jeigu k=10, tai maksimalinė liekana z gali sudaryti paklaidą ų pagal formulą:

$$ų = z:N \quad . \quad . \quad . \quad . \quad . \quad . \quad : \quad . \quad . \quad . \quad . \quad . \quad . \quad (1$$

1) jeigu N=109 tai z:N=9:109=0,089

2) „ N=1009 „ z:N=9:1009=0,0089

3) „ N=10009 „ z:N=9:10009=0,00089

Ir t. t.

Kaip matome, jei masė N nėra mažesnė kaip 10000, o atsitiktinių, kas dešimtasis imamų, reprezentantų skaičius n tesudaro tik 10%, maximalinė paklaida dėl liekanos paneigimo negali prašokti 0,089%. Jei gi koeficientą k sumažinti ligi 5 (stebėti 20%), tai maksimalinė arba vad. galimoji paklaida prie N = 1004 negali būti didesnė kaip 0,004 arba 0,4%. O kadangi statistikos stebimos masės gana dažnai turi ne tūkstančius, o šimtus tūkstančių ar net milijonus vienetų, tai galimoji paklaida dėl liekaninių dydžių tokiais atvejais pasidaro tiek nereikšminga, kad ją galima visai ir paneigti.

Taigi bendrajam masės vienetų skaičiui N nustatyti atsitiktinių reprezentantų metodas visai patikimas. Šiek tiek sunkiau yra statistinės masės skaitomųjų vienetų pažymius tokiu pat tikslumu apibrėžti. Bet ir šios rūšies sunkenybės savo esmėje yra tos pačios, kaip ir sunkenybės tikrajam st. vienetų skaičiui surasti.

Jei mes žinome, kad masėje N yra mums nežinomų dydžių daliniai a ir b, tai griežtai atsitiktinių reprezentantų sistemoj abidvi dalys turės vienodus šansus ir savuosius reprezentantus patiekti. Vadinasi, jeigu N masėje dalinys a=b, t. y. skaitomųjų vienetų su pažymiu a yra lygiai tiek, kiek ir b, tai ir atsitiktinių reprezentantų skaičiun galima tikėtis pateks iš abiejų (a ir b) dalinių maždaug po vienodą kiekį. Jei gi visoje masėje sk. vienetai su pažymiu a tesudarytų tik ⅓ dalį, tai ir reprezentantų skaičiuje, greičiausia, b pažymys žymiai dažniau pasirodytų, negu pažymys a. Klausimas tiktai, kaip didelė galėtų būti paklaida?

Teoretiškai imant, nėra negalimas daiktas, imant reprezentantu atsitiktinį dešimtąjį $\dot{x}_i$, vis palaikyti tik x_a arba tik x_b — jei iš viso kurio nors pažymio vienetų bendrasis skaičius yra nemažesnis, kaip bendrasis reprezentantų skaičius. Vadinasi, jeigu masėje $N=1000x_i$, $\Sigma x_a = 200$, ir $\Sigma x_b = 800$, o $\Sigma\dot{x}_i = 100$ (reprezentatu imamas kiekvienas dešimtasis, taigi k=10), — tai savaime yra galimas daiktas pataikyti visą 100 reprezentantų

su pažymiu a ir nė vieno — su pažymiu b. Tačiau tikimybė tokiam įvykiui yra labai maža, nes:

jeigu $\Sigma x_a = 200$, kai $N = 1000$, tai $\Sigma x_b = N - \Sigma x_a = 1000 - 200 = 800$.

O jeigu taip, tai ir tikimybė p reprezentantus tik iš Σx_a atsitiktinai pataikyti — bus keturis kartus mažesnė, kaip tikimybė — juos tik iš Σx_b pataikyti. Nes tikimybė

$$p_a = \frac{\Sigma x_a}{N} = \frac{200}{1000} = \frac{1}{5}, \text{ kai q arba } p_b = \frac{800}{1000} = \frac{4}{5}.$$

Ir tikrai: jeigu iš 1000 loterijos bilietų, 800 būtų „pilni" ir tik 200 „tušti", tai visi sakytų, kad esama kurios nors apgavystės, jei, 100 bilietų ištraukus, nebūtų pataikyta nė vieno „pilno". Aplamai, kaip praktika rodo, bent kiek daugiau loterijos bilietų „ištraukus", „pilnųjų" ir „tuščiųjų" santykis būna artimas tikrenybei. O matematika turi net priemonių apibrėžti šios rūšies tikimybių laipsnį. Savo esmėje šios rūšies skaičiavimai, taigi ir jų patikimumas, yra tokie pat, kaip ir anksčiau parodytieji skaičiavimai (iš tam tikro reprezentantų skaičiaus n masėje bendrajam vienetų skaičiui N surasti). Bet apie tai smulkiau ir nuodugniau galima rasti kituose skyriuose.

10. Statistinio stebėjimo organizavimas.

a) Bendrosios žinios.

Statistiškai prasmingas stebėjimas, t. y. statistinės medžiagos rinkimas ir reikalą atitinkančio metodo pasirinkimas, tik tuomet teįmanomas, kai visai aiškiai apsibrėžiama statistinio stebėjimo tikslas ir apimtis. Visų pirma, tenka išaiškinti, ar numatytam tikslui nepakanka statistinės medžiagos, kuri, galimas daiktas, jau kuriame nors statistikos šaltinyje yra paskelbta. O atsitinka, kad daromi statistikos surašinėjimai bet kuriam specialiniam reikalui ir tada, kai statistikos medžiaga jau yra kur nors paskelbta. Išaiškinus, kad reikalingos statistinės medžiagos niekur paskelbta nėra, tenka pirmiau, negu organizuoti statistikos stebėjimą, dar ištirti ir nustatyti, ar nėra kurioj nors įstaigoj reikalingos statistinės medžiagos pakankamai susikaupusios, tik dar neišnaudotos arba nepaskelbtos. Pav., mokesčių ir darbo inspekcijos arba teisingumo, susisiekimo, švietimo ir kitose ministerijose bei įstaigose beveik visuomet yra įvairiausios statistinės medžiagos. Ši medžiaga kai kada net ir sugrupuojama ir tų įstaigų reikalams naudojama, nors šiaip niekur neskelbiama. Kai kada šitokia medžiaga yra

visai tinkama ir betarpiškam statistikos metodais naudojimui, nedarant jokio naujo statistikos surašinėjimo. O net jeigu be statistinio stebėjimo nebūtų galima apsieiti, tai vis dėlto susipažinti su esamąja medžiaga yra naudinga, kad būtų galima nusistatyti, ar yra reikalinga organizuoti visai naujai betarpišką, t. y. tiesioginį statistikos stebėjimą, ar panaudoti statistikos reikalui esamą medžiagą netiesioginiu statistikos stebėjimo metodu. Tik galutinai visus šituos klausimus išaiškinus, galima pradėti ir patį statistikos stebėjimą organizuoti.

b) Surašomojo lakšto paruošimas.

Bet kuriuo metodu ir bet kuria sistema daromas statistinis stebėjimas yra reikalingas tikslingai paruošto ir prasmingai sudaryto surašomojo lakšto. Pati surašomojo lakšto forma ir turinys pareis nuo pasirinkto statistinio stebėjimo metodo. Vadinasi, individualinis statistinio stebėjimo metodas kortelių sistema bus reikalingas vienos surašomojo lakšto formos, o tas pats indvidualinis statistikos stebėjimo metodas sąrašų sistema — kitokios. Dar kitokia surašomojo lakšto forma ir turinys turės būti, jei patsai statistinis stebėjimas daromas ne individualiniu, o kolektyviniu metodu, t. y. jeigu surašomuoju vienetu imamas ištisas skaitomųjų vienetų kolektyvas. Surašomojo lakšto forma nei turinys, tačiau, nepareina nuo to, ar statistinis stebėjimas yra tiesioginis ar netiesioginis ir ar jis daromas ištisiniu, t. y. visuotiniu, ar daliniu, t. y. reprezentantiniu, stastikos stebėjimo metodu.

c) Statistinės klasifikacijos planas.

Statistiškai prasmingą surašomąjį lakštą sudaryti daug lengviau, jeigu iš anksto yra aiškiai apibrėžtas ne tik statistinio stebėjimo uždavinys bei tikslas, bet ir statistinio vaizdavimo pagrindiniai bruožai arba statistinės klasifikacijos planas. Dėlto visuomet rekomanduojama, kad bereikalingai negaišinti laiko ir neaikvoti lėšų, statistinio stebėjimo organizaciją pradėti arba iš karto nuo statistikos lentelių (tabelių) projektų sudarymo, arba, projektuojant statistikos surašomąjį lakštą, tuo pačiu laiku projektuoti ir numatomos tolimesnės medžiagos lentelių sudarymą. Galimas daik-

tas, kad ir šiuo keliu einant vėliau, statistikos medžiagą surinkus, viena ar kita pavaizduotos klasifikacijos statistinė lentelė pasirodys ne visai tinkama ar net nereikalinga. Vis dėlto, surašomąjį lakšto projektą iš anksto derinant su statistinės klasifikacijos lentelių projektu, bus išvengta visos eilės organizacinio ir redakcinio pobūdžio trūkumų. O gali net atsitikti, kad, bederinant surašomąjį lakštą su numatomos klasifikacijos projektu, paaiškės ir vieno ar kito statistikos stebėjimo metodo pranašumas.

d) Statistiškai prasmingo surašomojo lakšto ypatybės.

Svarbiausioji statistiškai prasmingo surašomojo lakšto ypatybė, tai jo t e c h n i k i n i s p a v a i z d u m a s i r s ą v o k i n i s g r i e ž t u m a s. Yra reikalinga, kad statistikos stebimųjų duomenų surašinėtojas, surašomąjį lakštą į rankas paėmęs, iš karto visą surašomųjų žinių kompleksą prieš akis turėtų arba čia pat jas peržvelgti ir numatyti galėtų. O iš to seka ir tam tikri statistiškai prasmingam surašomajam lakštui statomieji reikalavimai:

1) yra geriau, kad, prie kitų lygių sąlygų, surašomasis lakštas yra mažesnis ir jame pastatyti klausimai kiek galint ryškiau (didesnėmis raidėmis, retesnėmis eilutėmis) apibrėžti;

2) yra geriau, jei, prie kitų lygių sąlygų, visi klausimai yra surašyti vienoj surašomojo lakšto pusėj;

3) yra geriau, jei bet kurie su statistiniu stebėjimu susieti paaiškinimai yra čia pat prie statomų klausimų, negu toki paaiškinimai nukeliami į kitą lakšto pusę ar kur nors atskirai;

4) yra geriau, jei, prie kitų lygių sąlygų, surašomajam lakšte pastatytų klausimų yra mažiau, negu daugiau.

S ą v o k i n i s g r i e ž t u m a s statistiškai prasmingo surašomojo lakšto yra didelis palengvinimas ne tik surašinėtojams, bet ir būtinas, kad visas statistinis stebėjimas niekais nenueitų. Čia vėl reikalinga keletą būtinų sąlygų:

1) kiekviena surašomoji žymė turi būti griežtai nuo kitos atskirta, o įrašytini pažymiai geriausia ten pat tekste apibrėžti;

2) kiekvienas skaitomasis vienetas turi būti apibrėžtas o b j e k t y v i n i a i s pažymiais ir surašinėtojo lengvai išskiriamas;

3) surašomojo lakšto tekstas (žodinės skaitomojo vieneto pažymių sąvokos) turi būti aiškiai apibrėžta ir, be to, atsižvelgta galimų nevienodų tos pačios sąvokos atskiruose rajonuose apibrėžimų arba pavadinimų.

Šių sąlygų nevykdžius gali atsitikti nebepataisomų statistinio stebėjimo paklaidų[18].

e) Statistikos stebėjimo kontrolė.

Statistiškai prasmingas stebėjimas arba surašymas yra reikalingas kontrolės. Kontrolė turi apimti statistikos stebimą objektą, stebėjimo organizaciją ir patį stebėjimą.

1) Statistikos stebimo objekto kontrolė turi tikslo patikrinti:

a) ar surašomajam lakšte apibrėžtieji skaitomieji vienetai duodasi iš objektyvinių pažymių vienas nuo antro atskiriami,

b) ar numatytas stebėjimo metodas neturi kokių trūkumų, kurie kliudytų stebimą masę visapusiškai ir pilnai stebėti.

2) Statistinio stebėjimo organizacijos kontrolė turi parodyti:

a) ar surašomųjų lakštų forma ir turinys atitinka stebimo objekto surašytinas žymes bei pažymius,

b) ar stebėtojai bei surašinėtojai, taip pat tie, kurie stebėjimui reikalingas žinias turės tiekti, apibrėžtus ir statomus klausimus vienodai bei teisingai supranta,

c) ar į pastatytus per surašymą klausimus yra galima visai aiškius ir pilnus atsakymus duoti,

d) ar numatytas statistiniam stebėjimui (surašinėjimui) laikas yra patogus ir pakankamas.

3) Statistinio stebėjimo kontrolė turi parodyti, ar statistinis stebėjimas padarytas visiems numatytiems skaitomiesiems vienetams be paklaidų.

---

[18] Vokietijos statistikos praktikoj yra buvęs atsitikimas, kad Bavarijos ūkininkai klausimą apie vėjinius malūnus suprato kaip klausimą apie arpas. Ir statistinis stebėjimas „rado“ pas kiekvieną ūkininką po vieną „vėjinį malūną“, kai iš tikrųjų tie ūkininkai teturėjo tik arpas.

Pirmųjų dviejų punktų apibrėžti tikslai gali būti pasiekti tik per paruošiamąją kontrolę. Ši paruošiamoji kontrolė turi apimti tiek numatomą statistinio stebėjimo objektą (stebimą masę), tiek stebėjimo metodus bei organizaciją. Pav., darant gyventojų, įmonių ar kitokį statistinį stebėjimą yra visuomet pravartu iš anksto padaryti ir patikrinti stebimų objektų sąrašus, kurie yra (arba gali būti) kurioj nors administracijos įstaigoj (savivaldybėj, inspekcijoj ar kt.). Šios netiesioginiu keliu gautosios žinios įgalina jau iš anksto tikslingai apibrėžti surašomuosius rajonus, numatyti reikalingą kiekvienam rajonui surašinėtojų skaičių ir t. t. Bet kad patikrinti, kiek pasirinktas statistinio stebėjimo metodas yra tikslingas, gerai ar negerai sudaryti surašomieji lakštai, tikslingai ar netikslingai apibrėžti skaitomieji ir surašomieji vienetai ir jų pažymiai, kaip paruošti patys surašinėtojai — yra būtinai reikalinga vienu ar kitu metodu įvairiose vietovėse, atsižvelgiant jų skirtingų ypatybių, padaryti siauru mastu paruošiamosios kontrolės surašymą dar prieš galutinę surašomųjų lakštų redakciją ir prieš statistinio stebėjimo datų galutinį nustatymą. Praktikoje yra parenkami įvairiose vietose rajonai, iš surašinėtojų kadro paimami surašinėtojai ir duodama jiems padaryti bandomąjį surašymą tokia pat tvarka ir programa, kaip kad numatomas galutinis surašymas. Šito bandomojo surašymo patyrimais pasinaudojus, galima išvengti daugelio statistinio stebėjimo organizacijos klaidų ir netobulumų.

Pati statistinio stebėjimo kontrolė paprastai būna daugiau technikinio pobūdžio. Pav., jei statistinis stebėjimas yra organizuotas individualiniu metodu, tai iš surašinėtojų reikalaujama ne tik surašyti kiekvienam surašomajam vienetui korteles aiškiai apibrėžtame rajone (surašymo apylinkėse), bet ir suskirstyti surašomąsias korteles vienu ar kitu pažymiu ir, jas sudėjus į konvertą arba surišus, sudaryti savajai surašymo apylinkei paprastąją lentelę. Savaime suprantama, kad vienu ar kitu pažymiu sugrupuotų skaitomųjų vienetų bendrasis skaičius ir tiems vienetams surašytų pažymių skaičius surašomuose lakštuose turi sutapti. Jei nesutampa, reikia šį pirminės klasifikacijos darbą patikrinti.

11. Priedėliai.

a) Surašomasis lakštas individualiniam stebėjimui kortelių sistema.

PIRMASIS VISUOTINIS LIETUVOS GYVENTOJŲ SURAŠYMAS 1923 M.

Surašymo rajono Nr.....

ASMENS LAPELIS NR.....

Surašymo apylinkės Nr.....

Surašomas kiekvienam asmeniui surašymo dienoje.

Adresas gyvenančiam valsčiuje	**Adresas gyvenančiam mieste**
Apskritis	Miestas .....
Valsčius	Gatvė .....
Gyven. vietos pavadinimas:	Namų Nr..... Buto Nr.....

.....

Ūkio (sodybos) pavadinimas .....

..... sąraše Nr. .....

1) a) Pavardė .....
   b) Vardas .....
2) Lytis: vyrišk. (vyr.) ..... moter. (mot.) .....
3) Amžius: kiek metų .....; vaikams mažiau 1 metų: kiek mėn. ..... arba dienų .....
4) Šeimos padėtis: nevedęs (nev.) ..... netekėjus (net.) ..... vedęs (ved.) ..... ištekėjusi (išt.) ..... našlys-lė (našl.) .... persiskyręs-usi (prsk.) .....

Vaikams iki 16 metų nesurašoma

5) Kurios tikybos bažnyčioj arba kurios tikybinės organizacijos apeigomis krikštytas .....
6) Tautybė .....
7) Jei svetimšalis (ne Lietuvos), tai kurios valstybės pil. .....

8) Ar moka skaityti (taip, ne) ....., rašyti (taip, ne) .....

9) Ar visada gyvena tame valsčiuje ar mieste (taip, ne) ..... Jei ne, tai kur: valsčius ..... miestas ..... šalis .....

10) Gimimo vieta (čia pat, ne) ..... jei ne, tai kur: valsčius ..... miestas ..... šalis .....

11) Gyvenimo šaltiniai (iš ko gyvena):

A. Pagrindinis:	B. Pašaliniai: Pažymėti kokie
a) Įvardyjimas ir rūšis pramonės, įstaig. ar verslo, kur dabar dirba	
b) Verslo pavadinimas	
c) Padėtis versle: 1) savininkas, ... 2) nuomininkas, ... 3) šeimininkas - valdytojas, ... 4) darbininkas, ... 5) tarnautojas, ... 6) dirbantis šeimos narys	
d) Jei dabar neturi jokio darbo, pažymėti kurios profesijos (verslo)	
e) Jei darbuotojas žemės ūkyje ar turi savo žemės (taip, ne)	

12) Jei be verslo, tai iš ko gyvena .....

13) Fiziniai (kūno) trūkumai (aklas abiem akim, nebylys, kurčias, šlubas, kuprotas) ir sužalojimai-luošumai (be vienos ar be dviejų rankų, be vienos arba dviejų kojų) ....

14) Fizinių trūkumų ir sužalojimų priežastis (iš prigimimo, nelaimingo atsitikimo, ligos, karo) ....

1923 m.....mėn.....d.

Parašas: { davusio žinias ..... ar surašinėtojo ..... }

b. Surašomasis lakštas individualiniam st. stebėjimui sąrašų sistema.

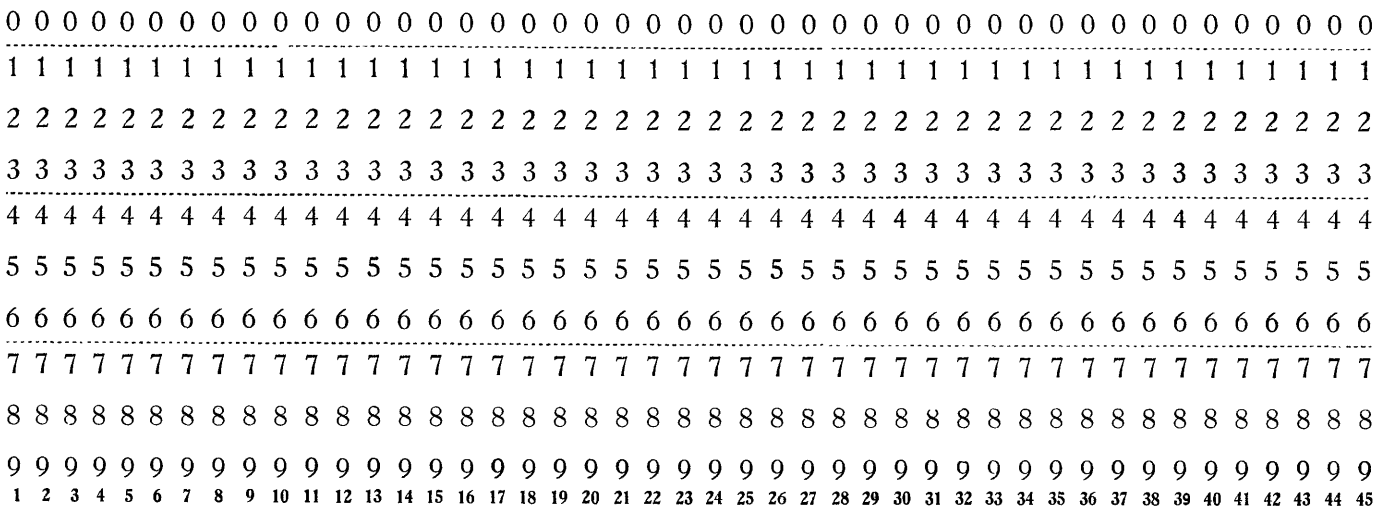
ŽYMĖJIMO LAPAS 1897 M. RUSIJOS VISUOTINIAJAM GYVENTOJŲ SURAŠYMUI DARYTI.

1	2	3	4	5	6	7	8	9	10	11	12	13		14	
Pavardė, vardas ir tėvo vardas arba vardai, jei keletas. Pažymėjimas, kas abiem akim aklas, kurčias, nebylys arba beprotis	Lytis: V — vyrams m — moterims	Kokiame santykyje šeimynos (ūkio) arba šeimos galvai.	Kiek turi metų arba mėnesių.	Nevedęs, vedęs, našlys arba atsiskyręs	Luomas, padetis (sostojenije) arba pašaukimas (zvanije).	Gimęs čia, o jeigu ne, tai kur? (Gubernija, apskr., miestas).	Prirašytas čia, o jeigu ne, tai kur? (asmenims, kurie privalo prisirašyti)	Kur paprastai gyvena — čia, o jeigu ne, tai kur? (Gub., apskr., miestas).	Nebuvimo, iškeliavimo arba laikinojo apsistojimo pažymėjimas.	Tikyba.	Gimtoji vieta.	Raštingumas		Užsiėmimas, amatas, verslas (promysel), vieta (dolžnost) arba tarnyba.	
												a. Ar moka skaityti.	b. Kur mokosi, mokėsi arba mokslą baigė.	a. Pagrindinis, t. y. toksai, kuris teikia didžiausią šaltinį gyventi.	b. 1. Pašalinis arba pagalbinis. 2. Santykis karo tarnybai.
1															1 / 2
2															1 / 2
3															1 / 2
4															1 / 2
5															1 / 2
6															1 / 2
7															1 / 2

c) Surašomojo lakšto pavyzdys kolektyviniam st. stebėjimui.

Nr.	Buto šeimininko pavardė	Bendras buto gyventojų skaičius	Lytis		Tautybė						Pilietybė		Šeimos padėtis				Amžius metais				
			Vyriš.	Moteriš.	Lietuvių	Žydų	Lenkų	Rusų	Vokiečių	Kiti	Lietuvos	Svetimšaliai	Vedę	Nevedę	Našliai	Išsiskyrę	Ligi 1 metų	1—2	2—3	3—4	Ir t. t.
1																					
2																					
3																					
4																					
5																					
6																					
7																					
8																					
9																					
10																					
11																					
12																					

d) Pavyzdys skaitomosios kortelės, naudojamos mechaniškam skaitomųjų žymių ir pažymių apdirbimui (klasifikacijai) Hollerith mašinų sistemoj.



TREČIOJI DALIS.

Statistikos vaizdavimas bei klasifikacija.

1. Statistinio vaizdavimo esmė ir uždaviniai.

Vienu ar kitu st. stebėjimo metodu gautoji ar surinktoji medžiaga yra tik žaliava. Ši žaliava, kad tiktų vartoti, yra reikalinga apdirbti, t. y. paruošti ir patiekti tokioj formoj, kad būtų st. metodais stebėtosios masės vaizdas matyti. Toksai vaizdavimas atliekamas st. duomenų grupavimu bei klasifikavimu.

Statistiškai prasmingas grupavimas bei klasifikavimas yra galimas tiktai prie sąlygos, jei statistinis stebėjimas yra atliktas visais atžvilgiais korektiškai, t. y. jei statistikos stebėtojai stebimos masės nė jos sudėtinių dalių bei charakteringų požymių neiškreipė. Dėl to yra pravartu prieš pradedant pagrindinį statistinės medžiagos apdirbimą, turimųjų duomenų patikimumą patikrinti — bent tiek, kiek šioj darbo stadijoj patikrinimas ir klaidų suradimas yra galimas. Tokio patikrinimo metodų ir priemonių yra ne viena, ir jau statistinės technikos uždavinys kiekvienam atskiram atvejui, ar konkretiškai paimtam uždaviniui, pasirinkti tinkamiausią tikrinimo metodą.

Kiekvieno grynai statistinio darbo uždavinys atvaizduoti masės arba masinio reiškinio charakteringas ypatybes. Tiesa, kai kada statistinis stebėjimas yra daromas iš anksto apibrėžtam kuriam nors reikalui arba tikslui (tikslinė statistika). Pav., iš anksto gali būti nustatyta, kad statistinis stebėjimas yra reikalingas tik mokyklinio amžiaus vaikų skaičiui surasti, krašto gynimui tinkamų vyrų kontingentui apibrėžti, mūrinių trobesių ar kurios nors vienos veislės ar amžiaus gyvulių skaičių nustatyti ir t. t. Tokiais ar panašiais atvejais iš statistikos stebėjimo surinktosios medžiagos, tą

medžiagą apdirbant, statistinių grupių bei klasių pažymiai ir jųjų skaičius savaime aprėžiamas statistikai pastatytojo uždavinio. Tačiau, kai statistinis stebėjimas yra daromas masės ar masinio reiškinio charakteringoms ypatybėms pažinti aplamai imant, tai statistinis vaizdavimas, t. y. grupavimas bei klasifikacija, privalo ne tik apdairumo, bet ir sistemos bei mokėjimo tą sistemą pritaikinti. Priešingu atveju statistikos vaizduojamos masės charakteringosios žymės ir jų pažymiai gali būti iškreipta ar net ir paslėpta.

2. Statistinės klasifikacijos rūšys ir jų metodologinis pagrindimas.

Paprastai statistinės metodologijos ir teorijos raštuose kalbama apie t r e j o p ą, t. y. trimis aspektais vykdomą, klasifikaciją:

1) statistikos duomenų grupavimas bei klasifikavimas e s m i n ė m i s (kilmės, rūšies etc.) žymėmis,
2) tų pačių duomenų klasifikacija p l o t o (arba erdvės) pažymiais ir
3) l a i k o aspektas nulemia statistinių grupių bei klasių sudarymą.

Trumpai kalbant, skiriamas esminis, plotinis ir laikinis (laiko pažymiais) grupavimas. Be to, kaip atskira statistinio vaizdavimo šaka su savomis problemomis yra k i e k y b i n ė klasifikacija arba d y d ž i o grupių sudarymas.

Tačiau neatrodo, kad šitokia statistinio vaizdavimo klasifikacija metodologiškai būtų galima pagrįsti. Taip pat nematyti, kad toki statistinės klasifikacijos apibrėžimai palengvintų statistinio vaizdavimo darbą, arba jį daugiau įprasmintų.

Jei mes gerai įsižiūrėsime bet kurį surašomąjį lakštą ir norėsime tame lakšte įrašytus pažymius klasifikuoti skirtingomis sistemomis, tai rasime, kad tokių sistemų galės būti ne daugiau kaip dvi: vienus pažymius, norėdami juos apibendrinti, gausime apibrėžti ž o d ž i a i s arba žodinėmis sąvokomis, kitus gi — skaičiais arba skaičių apibrėžiamomis sąvokomis. Pirmosios kategorijos priklauso tokios žymės bei pažymiai, k. a.: įvykio buvimo vieta, rūšis (lytis, veislė, kilmė, medžia-

gos ypatybės), paskirtis ir t. t. Sudarant statistines grupes arba tas grupes jungiant į klases, visas dėmesys turi būti kreipiamas tik į tai, kad objektyviniais duomenimis bei žodinėmis sąvokomis apibrėžtus pažymius (lytį, tautybę, vietovę ar kt.) išskirti ir jų nesukeisti, t. y. nesumaišyti. Kitokių statistiškai metodologinių problemų šios rūšies klasifikacijoj nėra, arba beveik nėra: statistikos darbininko svarbiausias ir vienintelis uždavinys — nepadaryti skaičiavimo arba sumavimo paklaidų.

Kitoks jau darbas klasifikuojant tuos pažymius, kurie surašomajame lakšte yra atžymėti skaičiais, k. a. kieno nors amžius, metai arba mėnesis, judėjimo statistikoj stebėjimo periodai, intervalai ir t. t., taip pat svorio, ilgio ar tūrio matai ir t. t. Visais šitais atvejais kurio nors skaitomojo vieneto priklausomybę vienai ar kitai grupei nulemia skaičius, t. y. k i e k i s: metų kiekis, mėnesių, dienų kiekis, tonų, kilogramų kiekis, metrų, kilometrų kiekis, litų, dolerių kiekis ir t. t. Žodžiu, bet kurios rūšies masės arba masinio reiškinio skaitomųjų vienetų grupinę priklausomybę nulemia ne kuris nors esminis pažymys arba kokybė, bet kiekybinis pažymys arba k i e k y b ė, t. y. kuris nors skaičius.

Dėl to, atrodo, nėra jokio pagrindo statistines grupes bei klases metodologiniu atžvilgiu skirstyti daugiau kaip į dvi sistemas: kokybinę ir kiekybinę. Visos tos žymės ir pažymiai, kuriais sudaromos grupės arba klasės, paprastai, rodo arba kurią nors materialinę bei formalinę ypatybę (kilmė, rūšis, buvimo arba veikimo sritis ir t. t.), arba kurios nors ypatybės vienu ar kitu (tūrio, svorio, erdvės, laiko) saiku matuojamą dydį. Gal tik plotinė klasifikacija kai kada gali atrodyti ne visai suderinama su kiekybinės klasifikacijos pagrindais. Bet ir čia tebus tik tariamas nesuderinamumas arba formalinis, o ne esminis. Kai statistikos duomenys grupuojami arba klasifikuojami administraciniais arba šiaip konvencionaliniais ploto vienetais, pav., valsčiais, apskritimis, valstybėmis arba kuriais nors kitais rajonais, tai čia visuomet turima reikalas su kokybine klasifikacija. Vieno ar kito skaitomojo vieneto vienai ar kitai geografiškai administracinei grupei priskyrimą nulemia vien tik to vieneto buvimo faktas iš anksto objektyviškai apibrėžtame ir žinomame plote — lygiai taip pat, kaip to

paties ar kito skaitomojo vieneto priskyrimas vienos ar kitos verslinės padėties, tikybos, tautybės ar kt. grupei nulemia atitinkamo verslo, tikybos, tautybės priklausymas. Šiais visais atvejais statistinė klasifikacija jokių naujų metodologinio pobūdžio problemų nekelia. Net tais atvejais, kai plotinės klasifikacijos pagrindan dedamas griežtai g e o g r a f i n i s kriterijus, t. y. kai masės skaitomieji vienetai grupuojami ne formališkai konvencionaliniais arba administraciniais vienetais (valsčiais, apskritimis ir t. t.), o geografiniu požiūriu (pav., atsižvelgiame tam tikrų rajonų miškingumo, drėgmės gausumo, jūros artumo ar upės ir t. t.), — visuomet turima reikalas su tos pačios rūšies problemomis, kaip ir grupuojant bet kuriais esminiais arba kokybiniais pažymiais. Erdvė arba plotas, dedamas statistinės klasifikacijos pagrindan, tik tada sudaro naujų, nuo kokybinės klasifikacijos skirtingų, metodologinio pobūdžio problemų, kai plotinio mato vienetais yra aprėžiamos skaitomųjų vienetų d y d ž i o grupės, t. y. kai vieno ar kito skaitomojo vieneto priskyrimą vienai ar kitai grupei nulemia p l o t i n i o mato pažymys (pav., ūkio didumas ha). Bet šiuo atveju ploto pažymys yra jau kiekybinis matas ir šios rūšies klasifikacija priklauso kiekybinės klasifikacijos.

3. Kokybinė klasifikacija ir jos uždaviniai.

a) Statistinis tikslingumas bei prasmingumas kokybinėj klasifikacijoj.

Kokybinės klasifikacijos uždavinys — turimus statistikos duomenis sugrupuoti taip, kad statistinės masės charakteringosios ypatybės būtų kiek galint ryškesnės. Šios rūšies grupavimas bei klasifikacija yra reikalinga tikslingo ir statistiškai prasmingo kokybinių grupių sudarymo ir jų statistikos l e n t e l ė s e arba t a b e l ė s e surašymo. Metodologinio pobūdžio problemos bei įvairios sunkenybės šios rūšies darbo stadijoj kyla iš to, kad vadinamas tikslingumas ir statistinis prasmingumas ne visuomet lengvai duodasi suderinami.

S t a t i s t i n i s t i k s l i n g u m a s st. medžiagą apdirbant gali būti įvairiai suprantamas bei aprėžiamas. Vienam reikalui gali būti reikalinga vienos rūšies kokybinės grupės, kitam — kitos; vienas gali interesuotis tik kuriuo nors vienu ar

nedaugeliu kokybinio pobūdžio pažymiu, o kitas bus reikalingas duomenų apie visus kurios nors esminės ypatybės (žymės) pažymius. Tačiau, jeigu būtų norima statistikos medžiagą patiekti sugrupuotą visais statistinio stebėjimo metu surašytais pažymiais, tai statistinis vaizdavimas lentelėse būtų gana sunkus ir ištaigingas darbas. Pav., 1923 m. gyventojų surašyme tautybės pažymių buvo atžymėta daugiau kaip 20. Taip pat po keliasdešimt pažymių buvo tikybos, pilietybės ir kt. Verslinės padėties pažymių buvo surašyta apie 1500. Vokietijoj gi per vieną statistinį stebėjimą verslinės padėties pažymių skaičius siekė apie 25.000. Visais šitais pažymiais statistikos duomenis sugrupuoti (sudaryti tiek kokybinių grupių, kiek yra pažymių) kad ir būtų techniškai įmanoma, bet užimtų be galo daug laiko. Statistinės gi lentelės pasidarytų tokios didelės, kad jose ir susivokti būtų nelengva. Vadinasi, visų arba daugybės detalių reikalaujanti t i k s l i n ė klasifikacija gali būti nesuderinama su statistinio pavaizdumo reikalavimais, o tuo pačiu — priešinga statistiniam prasmingumui.

S t a t i s t i n i s p r a s m i n g u m a s klasifikuojant statistikos medžiagą, visų pirma, reikalauja, kad sugrupuotoji statistikos medžiaga neiškreiptų charakteringųjų masės pažymių ir jų nenustelbtų. Be to, tuo pačiu žygiu reikalaujama, kad suklasifikuota statistikos medžiaga būtų pavaizdi ir praktiškai tinkama bet kuriam reikalui naudoti. Iš to seka pagrindinis reikalavimas, kad statistikos stebėjimo metu gautoji medžiaga pirmiau būtų suklasifikuota į tokias esminių pažymių grupes, kad vėliau kuriam nors specifiniam reikalui būtų galima jas jungti į kokybines klases. Vadinasi, kokybinė klasifikacija yra reikalinga, tarytum, dviejų etapų: 1) kiek galint didesnio kiekio, tačiau statistinio prasmingumo aprėžto, kokybinių grupių sudarymo ir jų lentelėse surašymo ir 2) jau daugiau tikslinių, taip pat kokybiniais pažymiais sudarytų, klasių ir jas atitinkančių statistikos lentelių.

Kiek pačių kokybiniais pažymiais sudarytų grupių ar klasių turėtų būti, tai čia jau pareina nuo statistinės medžiagos ir numatomų uždavinių bei tikslo. Pagrindinis reikalavimas ir čia, kaip ir visur statistikoj, kad sudarytosios kokybinės grupės bei klasės būtų statistiškai prasmingos ir neiškreiptų statistinės masės charakteringųjų ypatybių.

b) Statistiškai prasmingos klasifikacijos sąlygos.

Yra keletas pagrindinių sąlygų bei reikalavimų, kurios kokybinių grupių bei klasių statistinį prasmingumą gali nulemti.

1) Sudarant kokybines grupes arba klases ir duomenis vaizduojant statistikos lentelėje yra reikalinga, kad statistinių grupių arba klasių visuma sudarytų statistinį vienetą. Toks statistinis vienetas gali būti gaunamas tik tuo atveju, kai kokybinės grupės arba klasės sudaromos tik vienos kurios kokybinės žymės (ypatybės) pažymiais. Pav., kokybinė klasifikacija tautybės pažymiais gali turėti atskirų tautybių grupes (lietuvių, lenkų, rusų ir t. t.), klasifikacija tikybų pažymiais gali turėti atskirų tikybų (katalikų, protestantų, pravoslavų, izraelitų ir kt.) grupes ir t. t. Negalima į vieną lentelę jungti kokybines grupes iš dvejopos rūšies arba nevienodos kilmės pažymių. Pav., negalima vienoje st. lentelėje vienas grupes apibrėžti tautybės pažymiais, o kitas — pilietybės, tikybos ar kt. Šiam pagrindiniam reikalavimui Lietuvos statistikos kokybinėj klasifikacijoj yra, pav., prasižengta 1923 m. gyventojų surašymo duomenų klasifikacijos tautybės pažymiais. Ten į vieną lentelę buvo sujungtos atskirų tautybių grupės ir čia pat patiekta svetimšalių grupė, — nežiūrint į tai, kad toje svetimšalių grupėje, be abejonės, taip pat buvo atskirų tautybių žmonės.

2) Kokybinėj klasifikacijoj kiekviena grupė bei klasė turi būti aiškiai apibrėžta taip, kad jokiam atsitikime nebūtų galima jos sumaišyti arba sukeisti su kita grupe. Tačiau statistinis prasmingumas (statistinis pavaizdumas) taip pat reikalauja, kad kokybinių grupių skaičius nebūtų bereikalingai padidintas. Tam tikslui paprastai labai negausingos atskirų pažymių grupės jungiamos į vieną grupę, kuri jau apibrėžiama.

Pagrindinis statistikos reikalavimas yra tas, kad statistinę masę sudarančios griežtai apibrėžtų pažymių grupės apimtų bent du trečdaliu visos statistinės masės, o dar geriau — jei, perdaug neišplečiant grupių skaičiaus, pavyksta apibrėžtais pažymiais sudaryti kokybines grupes, kur vienetų bendroji suma

apimtų apie 95% visos statistinės masės. Šios rūšies reikalavimas, kaip toliau turėsime progos pamatyti, yra pagrįstas ir matematiškai tikimybinio skaičiavimo dėsniais (žiūr. skyrių apie st. dispersiją ir trijų sigmų taisyklę).

3) Vienais ar kitais principais statistinę masę klasifikuojant, be įvairių smulkmenų, kurios jungiamos į vieną kokybinę grupę paprastai su pažymiu „kiti" arba „kitoki", beveik visuomet atsiranda mažesnis ar didesnis skaitomųjų vienetų skaičius, kur dėl vienų ar kitų priežasčių iš visos surašomosios žymės pažymiu neatžymėta (darant statistinį stebėjimą užsimiršta įrašyti, negalima buvo sužinoti). Nėra tikslinga šios rūšies „nežinomųjų" grupę, kad ir nelabai ji būtų gausinga, jungti prie grupės „kitokių", t. y. įvairių atskiros neaprėžtųjų pažymiu smulkmenos. Statistinės klasifikacijos praktikoj tokiais atvejais elgiamasi dvejopai:

a) kai sudaromos ir smulkesnės kokybinės grupės, „nežinomojo" pažymio vienetai taip pat jungiami į atskirą grupę su apibrėžimu „nežinomi",

b) kai sudaromos kokybinės klasės griežtai apibrėžiant tik pačias gausingąsias kokybines grupes, „nežinomųjų" grupė panaikinama, jos narius prieš tai paskirstant proporcionaliai sulig lyginamuoju svoriu žinomųjų ir apibrėžtųjų pažymių grupėms.

Tik kai nežinomųjų grupė visai negausinga, nebus didelės nuodėmės, jei jau iš anksto tuos kelius vienetus su „nežinomu" pažymiu priskirsime kuriai nors vienai ar kelioms iš gausingiausiųjų grupių.

4. Kiekybinė klasifikacija (dydžio grupių bei klasių sudarymas).

a) Kiekybinės klasifikacijos tikslas ir bendrieji dėsniai.

Kiekybinė klasifikacija turi parodyti st. masės (masinio reiškinio) kiekybines ypatybes, t. y. masės skaitomųjų vienetų lyginamojo svorio, tūrio, ploto, apimties, aukščio, vertės, amžiaus ar kurios nors kitos kiekybės ypatybes. Šiam tikslui sudaromos vadinamos dydžio grupės bei klasės griežtai apibrėžtų kiekybių (svo-

rio, tūrio, ploto ir kt.) matais (kilogramais, litrais, metrais, mėnesiais ar kt.). Statistiškai metodologinio pobūdžio problemos kiekybinėj klasifikacijoj yra susietos tiek su kiekybinių grupių skaičiumi, tiek su jų pažymio (intervalų) aprėžimu. Ir tas problemas sudaro pirmoj eilėj vėl iš statistinio tikslingumo bei prasmingumo išeinančių reikalavimų susikirtimas ar bent nesutapimas.

Statistinis prasmingumas, visų pirma, reikalauja, kad kiekybinė klasifikacija kiek galint tiksliau atvaizduotų kiekybinį statistinės masės charakterį su jo ypatybėmis. Šitas, tačiau, reikalavimas kai kada gali nesutapti su statistinio tikslingumo reikalavimais. Šios rūšies kolizija ypač gali įvykti tais atvejais, kai statistikos duomenys ir jų kiekybinė klasifikacija yra reikalinga kuriam nors apibrėžtam reikalui, pav., kai interesuojamasi kurio nors iš anksto apibrėžto dydžio grupėmis arba klasėmis, o ne kiekviena bet kuria kiekybinio susigrupavimo smulkmena. Pav., švietimo politikos reikalui gyventojų statistikoj svarbiausios ir įdomiausios yra vadinamos mokyklinio amžiaus grupės, kai krašto apsaugos arba militarinės politikos reikalui — įdomiausios yra karui tinkamų vyriškių amžiaus grupės. Vėl kitoki reikalavimai statomi kiekybinėms amžiaus grupėms gyvybės draudimo interesantams, dar kitoki — vadinamos demografinės bei socialinės politikos reikalams. Ir t. t. Taip pat nevienodi ir ne tie patys reikalavimai statomi kiekybinei klasifikacijai ir kitos rūšies statistinėms masėms.

Šiuos čia minėtus ir neminėtus kiekybinei klasifikacijai statomus reikalavimus gali patenkinti tik tokia kiekybinė klasifikacija, kuri sugebės pavaizdžiai, t. y. be nereikalingų ar mažai vertingų smulkmenų sudaryti dydžio grupes ir jas patiekti statistinėj lentelėj taip, kad visos kiek ryškesnės ir svarbesnės kiekybinio pobūdžio grupės griežtai vienodu principu būtų aiškiai apibrėžtos. Tuo būdu sudaromos sąlygos, kur statistikos duomenų vartotojai arba naudotojai, susipažinę su masės kiekybinio pasiskaidymo charakteringomis ypatybėmis, savo specifiniam reikalui gali imti bei analizuoti, arba ir savaip klasifikuoti, tik tas kiekybines grupes, kurios jiems atrodo tinkamiausios arba reikalingiausios.

Vadinasi, pirminė statistinės medžiagos kiekybiniais pažymiais klasifikacija privalo atvaizduoti masės skaitomųjų vie-

netų kiekybinį, t. y. sulig jų lyginamuoju svoriu bei dydžiu sudarytą pasiskaidymą. Iš to seka, kad pirminė kiekybinė klasifikacija yra reikalinga tokių metodų, kurie įgalintų visai bešališkai (objektyviniais pažymiais) suskaidyti statistinę masę į dydžio grupes su griežtai apibrėžtais tų grupių intervalais. Iš įvairių galimų metodų pirmenybė tenka atiduoti tokiam metodui, kuris mažiausiai vietos palieka klasifikuotojo nuožiūrai ar subjektyviniam nusistatymui, ir daugiausia leidžia pasireikšti didžiųjų skaičių dėsniui. Tokį kiekybinį klasifikacijos bešališkumą garantuoja kai kurie statistiškai matematinės klasifikacijos metodai.

b) Statistinės klasifikacijos pagrindai.

Statistiškai matematinės klasifikacijos pagrindinis reikalavimas yra tas, kad kiekvienas statistinės masės vienetas traktuojamas vienodai, ir jo priklausomybę vienai ar kitai grupei arba klasei nulemia vien klasifikacijos pagrindan padėtieji kiekybiniai pažymiai. Tas vienodas visų masės vienetų traktavimas, tačiau, gali būti garantuotas tik prie sąlygos, kai kiekybinėms grupėms apibrėžtieji intervalai perdėm yra tie patys arba vienodi — išskiriant, žinoma, intervalus pačių kraštutinių kiekybinių grupių arba klasių, kai intervalo viena riba palieka neaprėžta. Pav., jei imti gyventojų statistikos duomenis ir kiekybinės klasifikacijos pagrindan padėti atskirų žmonių amžių, tai griežtas matematinio metodo pritaikymas pareikalaus iš anksto apibrėžti visoms amžiaus grupėms intervalo ribas. Statistiškai metodologiniu atžvilgiu nėra svarbu, ar gyventojų amžiaus pažymiai turės metinius, trimetinius, penkmetinius, dešimties metų ar kitokius intervalus. Svarbu, tačiau, kad vienai amžiaus grupei apibrėžtas intervalas būtų išlaikytas visose gyventojų amžiaus grupėse nuo jauniausio kūdikio ligi seniausiojo amžiaus žmogaus. Jeigu kiekybiniais pažymiais būtų klasifikuojami žemės ūkiai, klasifikacijos pagrindan padėjus vieno ūkio priklausantį žemės plotą, tai vienai dydžio grupei paimtas intervalas (tarsime — vieno, penkių, dešimties ha ploto ribose) privalo būti išlaikytas perdėm visai ūkių masei. Panašiu principu privalo būti klasi-

fikuojami ir duomenys vertės arba kainos, tūrio, svorio ar kt. mato vienetai.

Šiuo metodu sudarius dydžio grupes iš karto bus matyti, kurio, būtent, dydžio skaitomųjų vienetų yra daugiausia, kurių mažiau ir kurių visai mažai. Taip pat kiekvienai tokiai kiekybinei (dydžio) grupei tenkantis skaitomųjų vienetų skaičius ir pačių kiekybinių grupių pasiskaidymas parodo, ar visa statistinė masė yra susikaupusi apie kurį nors vieną aiškiai prasikišantį dydį, ar, gal būt, tyrinėjamoj masėj yra daugiau negu vienas dydžių bei jų grupių susikaupimas. Šitaip suklasifikuotos kiekybiniais pažymiais statistikos duomenų vartotojas jau lengvai sugebės ne tik išskirti statistinės masės dydžių susikaupimus, bet ir panaudoti tuos susikaupimus statistiškai prasmingam analiziui. Viena, tačiau, sąlyga šitokioj klasifikacijoj yra būtina: pirminėj kiekybinėj klasifikacijoj dydžio grupėms intervalai privalo būti neperdideli, kad negalėtų pasislėpti arba iškrypti statistinės masės natūrinio susikaupimo charakteringesnės, gyvenimiškai pagrįstos, ypatybės.

Vadinasi, kiekybinėj klasifikacijoj imamieji dydžio grupei visi intervalai turi atitikti ne tik matematinės klasifikacijos statomą pagrindinį reikalavimą (vienodi intervalai), bet ir turi būti suderinti su statistinės masės dalinių (dydžio grupių) charakteringomis ypatybėmis. O gali atsitikti, kad vienoje statistinės masės dalyje kurios nors rūšies charakteringos ypatybės pasireiškia griežčiau, ne kaip kitoje. Pav., jei paimsime gyventojų amžiaus grupes ir panagrinėsime jų dalyvavimą įvairiose socialinio gyvenimo srityse, tai tas dalyvavimas bus toli gražu ne vienodas. Štai, kriminalinės statistikos duomenys rodo, kad vaikai ir mažamečiai, taip pat senesniojo amžiaus žmonės, bendroje nusikaltėlių masėje sudaro labai mažą koeficientą. O amžiaus grupės maždaug nuo 20 ligi 40 metų nusikaltėlių masėje atstovaujami labai gausingai, nors ir ne vienodai. Taip pat atskirose nusikaltėlių kategorijose vienos grupės dalyvauja daugiau, kitos — mažiau. Suskirsčius visą statistinę medžiagą amžiaus grupėmis vienodais intervalais, sakysim, kiekvienai dydžio grupei duodant 10 metų intervalus, mes gausime iš karto kiekybines grupes, iš kurių kiekvienoj, gal būt, tik maža da-

lis, ir tai dar nepakaltinama, tedalyvauja bendroje nusikaltėlių masėje (tokios grupės pavyzdys bus pirmojo dešimtmečio amžius). Iš kitos pusės mes turėsime dydžio grupių, kur bus suplakta tokios masės, kurios savo dalyvavimu nusikaltėlių masėje arba atskirose nusikaltimų rūšyse dalyvauja labai nevienodai (pav., nuo 10 ligi 20 metų).

Arba imkime žmonių mirtingumą. Stebint mirusiųjų amžių ir jų santykį su atitinkamo amžiaus masės daliniais galima konstatuoti, kad, pavyzdžiui, mirtingumas atskirų metų amžiaus grupėse maždaug nuo 5 ligi 50 ar net 60 metų nedaug kuo tesiskiria. Vadinasi mirtingumo statistikai nieko nekenktų, jei būtų sudarytos amžiaus grupės ir žymiai didesniais, ne kaip dešimties metų, intervalais. Tačiau padėtis iš esmės pasikeičia, jei žiūrėti mirtingumo ne gyventojų masės nuo 5 ligi 60 metų, o, sakysime, iš vienos pusės ligi 5, o iš kitos pusės nuo 5 ligi 10 metų arba senesnių kaip 60 metų. Čia mirtingumo koeficientai jau ir metinėms grupėms gana daug skiriasi. O jei paimti jauniausiąją amžiaus grupę (kūdikius ligi 1 metų), tai rasime, kad ši amžiaus grupė bendroje mirusiųjų masėj ne tik gausingiausia dalyvauja, bet ir pačioje grupėje yra dar labai didelių skirtumų, jei tik ją suskaidyti mažesniais, sakysime, 1 mėnesio intervalais.

Iš patiektųjų pavyzdžių matyti, kad kiekybinė klasifikacija statistiškai matematiniu metodu pirminėms dydžio grupėms reikalinga tokių intervalų, kurie įgalintų pastebėti ir išskirti masės kiekybinio pobūdžio charakteringas ypatybes net ir smulkiuose daliniuose, jei tik šie daliniai, suskaidžius dar smulkesniais intervalais, pasirodo labai nevienodi. Tačiau ši aplinkybė iškelia techniškai metodologinio pobūdžio problemų. Visų pirma, šis reikalavimas, praktikoj vykdomas, veda ta kryptimi, kad pirminės klasifikacijos stadijoj dydžio grupės yra reikalingos kiek galint mažesnių intervalų. Bet juo bus mažesni intervalai, tuo daugiau bus ir kiekybinių grupių. Pav., turint galvoje, kad yra žmonių, turinčių ir daugiau kaip 100 m., jau metiniai dydžių grupių intervalai pareikalaus nemažiau kaip 100 skilčių. O jei panorėti kiekybinę klasifikaciją pritaikinti, pav., ir pirmamečių kūdikių atskirų mėnesių mirtingumo skirtumams, tai tektų imti ne metinius, o mėnesinius amžiaus grupėms intervalus. Bet čia amžiaus grupių skaičius tuojau

išaugtų ligi 1200, — nežiūrint į tai, kad aukštesnėse grupėse ištisos skiltys turėtų būti paliktos tuščios dėl skaitomųjų vienetų stokos.

Šios rūšies sunkenybės statistikos praktikoj gali būti nugalimos įvairiais būdais. Pav., jei pirminėj klasifikacijoj mažiau ar daugiau įprastais kurio nors mato vienetų intervalais susidaro labai gausingos dydžio grupės ne st. masės centre (apie vidurinį dydį arba vidurkį), o kuriame nors krašte arba gale, tai yra pažymys, kad šios rūšies masėje esama statistiškai savarankiškų dalinių, kurie su visa mase gali būti tik formaliai susieti (pav., seniai ir vaikai, dvarininkai ir mažažemiai ir t. t.). Šie daliniai tuo būdu yra reikalingi ir atitinkamų intervalų kiekybinėms grupėms sudaryti, kad būtų galima ir tų dalinių charakteringas ypatybes pamatyti bei atvaizduoti. Paprastai toki statistinės masės pakraščiuose susikaupę daliniai yra reikalingi arba dar daugiau išskaidyti smulkesniais intervalais, arba — jau stambesnių intervalų. Statistiškai metodologiniu atžvilgiu toki daliniai yra jau savarankiški arba autonomiški statistinės masės vienetai (pav., pirmamečiai kūdikiai skaidomi ne metinėmis, o mėnesinėmis amžiaus grupėmis, smulkesnieji ūkiai po 1 ha, jei visa ūkių masė suskirstyta penkių ar dešimties ha intervalais; iš kitos pusės stambesniųjų ūkių grupės skiriamos tik 20, 50 ar net 100 ha intervalais).

Iš kitos pusės, kad išvengt tuščiavidurių kiekybinių grupių tiems atvejams, kai kuris nors masės vienetas, arba mažas tų vienetų skaičius, labai jau išsiskiria iš masės, galima sudoroti taip, kad šituos per daug išsiskyrusius vienetus prijungiama prie pačios kraštutinės dydžio grupės su negausingiausiu vienetų skaičiumi. Šiuo būdu bus patenkinti ir matematinės klasifikacijos reikalavimai, ir išvengta bereikalingo kiekybinių grupių skaičiaus plėtimo. Šitaip sudarytos pirminės dydžio grupės įgalina visą medžiagą ir jos charakteringas ypatybes pažinti, o pažinus — reikale ir aukštesnės kategorijos dydžio klases sudaryti.

c) Aukštesniojo laipsnio kiekybinės klasės.

Aukštesniojo laipsnio kiekybinių klasių sudarymas yra reikalingas ypatingo stropumo. Statis-

tiškai metodologiniu atžvilgiu, aukštesnės kategorijos dydžio grupėms perdėm vienodi intervalai nėra reikalingi. Atvirkščiai, aukštesniojo laipsnio dydžio klasėms jų intervalai turi būti pritaikinti statistinės masės natūrinio pasiskaidymo charakteringoms ypatybėms ir, būtent, taip, kad kiekviena aukštesnio laipsnio dydžio klasė kiek galint geriau atvaizduotų statistinės masės vienetų dispersijos (susikaupimo bei pasiskaidymo) ypatybes. Pav., jei statistinės masės vienetai yra daugiau susikaupę apie patį vidurinį arba vad. centrinį dydį (medianą), tai turi būti sudaryta ir centrinė dydžio klasė su taip aprėžtu intervalu, kad šitoje dydžio klasėje patsai centrinis dydis (medianas) atitiktų intervalo vidurį[19]. Taip pat ir kitos šalutinės arba antraeilės, dydžio klasės privalo tokių ir taip aprėžtų intervalų, kad kiekvienoje dydžio klasėje didžiausias vienetų skaičius patektų intervalo centran. Jeigu be susikaupimo apie centrinį masės vienetą yra dar didesnių arba charakteringesnių susikaupimų kuriame nors masės sektoriuje (pasirodo labai gausingos smulkesniųjų arba stambesniųjų vienetų grupės), tai aukštesnės kategorijos klasifikacija turi tų susikaupimų atsižvelgti ir tuos susikaupimus atitinkančias grupes ypatingu intervalo aprėžimu išskirti. Kriterijus ir patys statistiniai dydžiai tokiems atvejams galima surasti atitinkamomis matematinio analizio priemonėmis, apie kurias plačiau bus sekančioje dalyje paaiškinta.

d) Kiekybinė klasifikacija Galton'o metodu.

Tam tikrais atvejais pirminė kiekybinė klasifikacija gali būti prasminga ir vadinamuoju G a l t o n'o metodu. Vad. Galton'o[20] metodo kiekybinės klasifikacijos esmė yra ta, kad visa statistinė masė, išrikiavus ją sulig vyresniškumu, kiekybiniais pažymiais suskaidoma į iš anksto aprėžtą grupių skaičių, kiekvienai grupei duodant vienodą skaitomųjų vienetų kiekį. Techniškai imant, statistinės masės vienetų bendroji suma yra da-

[19] Pav., jei dydžio klasės intervalas aprėžiamas 10—30 ha plotu, tai šios klasės centrinis dydis (medianas) privalėtų turėti apie 20 ha.

[20] F. Galton'as (1822—1911) laikomas eugenikos mokslo kūrėju. Jo tyrinėjimai matematinio analizio metodais daugiausia lietė paveldėjimo klausimus.

linama numatytųjų grupių skaičiumi, ir dalmuo parodo masės vienetų skaičių, tenkantį kiekvienai dydžio grupei. Jei, pav., visą statistinę masę, susidedančią iš N vienetų, norima suskaidyti į 10 kiekybinių grupių, tai daloma dešimtimi (N:10), jei į šimtą — tai šimtu (N:100), ir t. t. Kiekvienos tokios grupės intervalas aprėžiamas pačių kraštutinių tokios grupės skaitomųjų vienetų kiekybiniu pažymiu.

Šio metodo pagrindinė ypatybė yra ta, kad statistinės masės vienetų didžiausius susikaupimus — o tuo pačiu ir dažniausiojo dydžio vietą, rodo grupės su mažiausiais intervalais, mažiausius gi susikaupimus (rečiausią pasiskaidymą) rodo grupės su didžiausiais intervalais. Savaime suprantama, kad šis metodas tegali parodyti tik apytikrius masės vienetų susikaupimų centrus. Nes kai kiekvienos atskiros kiekybinės grupės intervalo ribas nulemia atsitiktinai paimtas statistinės masės vienetas (dalybos keliu a priori kraštutiniai skaitomieji vienetai aprėžiami), tai gali pasitaikyti, kad bet kurioj kiekybinėj grupėj centrinis dydis patenka ne į vidurį, o į bet kurį vieną intervalo kraštą. Šiuo būdu sudarant dydžio grupes, visai yra galima, kad statistinėje masėje natūrinis vienetų susikaupimas bus dirbtinai perskirtas pačiam susikaupimo centre. Dėl šios priežasties Galton'o metodas kiekybinėj klasifikacijoj yra ir gali būti su nauda naudojamas tik patyrusio statistiko ir tik specialiniams statistinio analizio uždaviniams spręsti.

e) Kumuliatyvinis metodas kiekybinėj klasifikacijoj.

Tais atvejais, kai kiekybinė klasifikacija turi uždavinį parodyti stebimos masės ar masinio reiškinio vienos kurios dalies santykį su visa mase, arba jos lyginamąjį svorį visoje masėje, pirminės klasifikacijos gautosios dydžio grupės, o kai kada iš karto ir natūrinės grupės, vad. kumuliatyviniu metodu jungiamos į atitinkamas dydžio klases. Jeigu statistinės klasifikacijos uždavinys yra parodyti tik kurio nors vieno pažymio lyginamąjį svorį (pav., mokyklinio amžiaus vaikus, darbingojo amžiaus gyventojus, karinės prievolės privalančius vyrus, tam tikrų mokestinių pajamų kategorijoms priklausančius asmenis ar kt.), tai iš viso tesudaromos tik dvi kiekybinės klasės. Iš šių klasių viena apima specifiniai atvaizduotą statistinės masės

dalį (pav., tarp 7 ir 12 m. mokyklinio amžiaus vaikai, 15 ar 18 ligi 60 ar 65 m. gyventojai, nuo 20 ligi 25 m. vyriškiai ir t. t.), o kita — visus kitus, kurie šios specifinės kategorijos nepriklauso. Be to, kumuliatyvinis metodas dydžio grupėms gali būti taikomas ir ne tikslinei (specifinei) klasifikacijai su griežtai aprėžtu uždaviniu, o ir šiaip kaip statistinio vaizdavimo priemonė statistinėms masėms jų tam tikroms ypatybėms aiškiau apibrėžti.

Aplamai, kiekybinės klasifikacijos bendrieji dėsniai yra tie patys kaip ir kokybinėj klasifikacijoj. Vadinasi, ir kiekybinėj klasifikacijoj griežtai aprėžtais intervalais dydžio grupės turėtų sudaryti ne mažiau kaip ⅔ visos statistinės masės, o dar geriau, jei griežtai aprėžtais intervalais dydžio grupės apimtų bent apie 95% visos masės (dviejų kvadratingųjų nukrypimų ribos prie normalinės dispersijos). Taip pat ir kiekybinėj klasifikacijoj nežinomo dydžio skaitomieji vienetai privalo būti išskirti, arba proporcionaliai paskaidyti tarp dydžio grupių aprėžtais intervalais. Savaime suprantama, kad kiekybinių grupių arba klasių sudėtiniai skaičiai turi atitikti bendrąjį statistinės masės skaitomųjų vienetų skaičių.

B. Statistikos klasifikacijos sistemos.

5. Statistikos lentelės ir terminai.

a) Statistikos lentelių sudarymas ir kai kurie st. technikos terminai.

Vienais ar kitais pažymiais suskirstytos masės arba masinių reiškinių daliniai (kokybinės arba kiekybinės grupės bei klasės) statistinio pavaizdumo sumetimais jungiami į statistikos lenteles arba tabeles. Statistikos lentelių sudarymo techniniai klausimai, taip pat svarbesnieji terminai šiandien jau yra nustatyti ir oficiališkai visiems rekomenduoti Tarptautinio Statistikos Instituto 23-joj sesijoj 1937 m. Atėnuose. Tos visuotiniai priimtos rekomendacijos apima šiuos klausimus:

1. Statistinė lentelė. Ją sudaro: a) antgalvis, b) skiltys, c) prieskiltis, d) skaičių eilutė ir e) lastelė.

a) Antgalvis dedamas st. lentelės viršuje, skiltyse duodamų skaičių turiniui pažymėti.

b) Skiltys eina statmeniškai po antgalviu ir jose patiekiami lentelės skaičiai.

c) Prieskiltis užima kairiąją st. lentelės dalį, ir čia pažymimas skaičių eilutės turinys.

d) Skaičių eilutė eina horizontaliai per visas skiltis.

e) Lastelė — tai vieta, kur skiltyje įrašomas eilutės skaičius arba, jo nesant, atitinkamas ženklas.

2. Kiekviena statistikos lentelė privalo būti taip sudaryta, kad kiekvienas jos skaičius būtų logiškam sąryšy su kitais lentelės skaičiais ir suprantamas iš pačios lentelės antgalvyje ir prieskiltyje įrašytų pažymėjimų bei ten pat atžymėtų pastabų.

Kiekviena lentelė privalo turėti skaičių arba, jo nesant, atitinkamą ženklą.

Jei leidinyje yra daug st. lentelių, kiekviena iš jų turi turėti savo turiniui atitinkantį pavadinimą ir eilės numerį.

Tarptautinis Statistikos Institutas rekomenduoja net ir st. lentelių normalinius formatus.

3. Normaliniai formatai st. leidiniams (biuleteniams, metraščiams ir t. t.) rekomenduojama šie (milimetrais):

420×594	297×420	210×297	148×210	105×148
		74×105		

Specialiniai formatai knygoms:

500×707	353×500	250×353	176×250	125×176
		88×125		

4. Sudarant st. lenteles reikia vengti skirtingų, taip pat nepaprastų arba labai retai vartojamų matų vienetų. Kiekvienu atveju st. lentelėj vartojami mato vienetai turi būti griežtai ir aiškiai apibrėžti.

Visi skaičiai, kurie turi daugiau kaip tris skaitmenis (išskiriant skaičius metams žymėti), privalo būti atskirti: arba tar-

peliu (pav., 2 345 895), arba kableliu (2,345,895), arba tašku (2.345.895).

5. Skaičių šriftas st. lentelėse rekomenduojama toks:

a) storas (juodas) šriftas sumoms žymėti,
b) kursyvas — dalmeniniams santykio skaičiams,
c) petitas — skaičiams, kurie lentelėje nesumuojami arba kurie iš kitų išskiriami.

Rankraštyje (ranka arba mašinėle rašytame) sumų skaičiai du kartu pabraukiami, dalmeniniai santykio skaičiai (kursyvas) vieną kartą pabraukiami ir nesumuojamieji skaičiai arba iš kitų išskirtieji, skaičiai (petitas) pabraukiami punktyru.

Šriftas privalo būti parinktas toks, kad lastelėse skaičiai paliktų ryškesni, t. y. per daug nesusiglaustų. Jei skiltys yra ilgos, yra rekomenduojama kas penktąją eilutę atskirti mažu tarpeliu, o kas dešimtąją — didesniu baltu tarpu (interlinijomis).

Jei vienoje st. lentelėje patiekiami duomenys daugiau kaip vienos žymės pažymiais sugrupuoti, tai vienos žymės pažymius apimanti skilčių grupė, kuri sudaro visumą, privalo nuo kitos skilčių grupės būti atskirta storais, pusiau storais arba dvigubais brūkšniais. Reikalinga vengti skersinių st. lentelių, taip pat lentelių su antgalvio pažymiais prieskilčio vietose.

6. St. lentelių numeracija. Jei st. lentelių yra daug skilčių, rekomenduojama kiekvieną skiltį atžymėti eilės numeriu arba abėcėlės raidėmis.

Jei st. lentelė viename puslapyje neišsitenka, tai eilučių tekstas prieskiltyje privalo būti pakartotas antrojo puslapio paskutinėje skiltyje. Rekomenduojama taip pat lentelės skaičių eilutes pažymėti eilės numeriais prieskilčio kairėje pusėje (prieš eilučių turinį apibrėžiamąjį tekstą). Ši lentelės eilučių numeracija gali būti panaudota tais atvejais, kai lentelė užima du puslapius ir kai prieskiltyje eilučių tekstas dėl vienų ar kitų priežasčių negali būti pakartotas antrame puslapyje. Tokiu atveju eilučių eilės numeriai pakartojami antrojo puslapio paskutinėje skiltyje, kuri gali būti siaura ir mažai vietos užimanti. Taip pat galima praktikuoti eilučių numeraciją pir-

mojo puslapio paskutinėje skiltyje ir antrojo puslapio pirmoje skiltyje tuo būdu, kad abiejuose puslapiuose esančios lentelės eilučių numeriai eitų gretimai esančiose skiltyse.

Statistikos lentelėse, kurios užima daugelį puslapių ir vienu atsivertimu visos lentelės matyti negalima, antgalviniai skilčių įrašai privalo būti pakartoti kiekviename puslapyje. Taip pat pastabos, kurios paprastai rašomos išnašose, privalo būti pakartotos kiekviename puslapyje visais tais atvejais, kai lentelė užima daugelį puslapių ir kai pastabos liečia visą lentelę.

7. Sudarant st. lentelę, kiekviena jos lastelė turi būti kuo nors užpildyta. Naturalu, žinoma, įrašyt statistinės eilutės atitinkantį skaičių. Tačiau gali atsitikti, kad vienai ar kitai ląstelei reikiamo skaičiaus nėra ar tai dėl to, kad kuris nors pažymys duotoje lastelėje visai neatstovaujamas, ar dėl to, kad reikalingų duomenų neturima, ar, pagaliau, dėl to, kad turimų duomenų yra permaža, kad lentelėje apibrėžtą eilutės minimalinį skaičių sudarytų. Dėlto yra sutarta ir rekomenduojama: — B r ū k š n e l į rašyti į atitinkančią lastelę tais atvejais, kai klasifikacijoj numatytam pažymiui duomenų, padarius statistinį stebėjimą, negauta (pav., tarptautinį reiškinį apimančioje lentelėj Lietuvai tenkančioje st. eilutėje šį ženklą tektų įrašyti lastelėje, kuri turėtų rodyti kupranugarių skaičių, nes kupranugarių Lietuvoje būti negali; taip pat tektų brūkšnelį įrašyti lastelėje žąsų eksporto lentelėje tais mėnesiais, kai žąsys neeksportuojamos).

. T a š k a s įrašomas į lastelę tais atvejais, kai reikiamų st. duomenų neturima arba jų negauta, o yra žinoma, jog atitinkančių skaitomųjų vienetų arba įvykių esama (pav., gali atsitikti, kad iš kurios nors vietovės arba kuriam nors skaitomajam periodui laiku negauta duomenų apie gimimus, mirimus ar kt.).

0. N u l i s įrašomas į lastelę tais atvejais, kai turimųjų skaitmeninių duomenų yra permaža, kad būtų galima į lastelę įrašyti reikalingas sveikas racionalus skaičius. Taip pat nuliumi atžymima dešimtainės trupmenos tais atvejais, kai lentelėje kitose lastelėse žymima tik sveikais vienetais.

Tokius bendrojo pobūdžio nuostatus rekomenduoja Tarptautinis Statistikos Institutas. Prie jų galime pridėti dar kai kurias kitas techniškai metodologinio pobūdžio pastabas.

b) Intervalo apibrėžimas kiekybinėj klasifikacijoj.

Kiekybinių grupių bei klasių apibrėžimas gali būti vykdomas skiriamuoju ir jungiamuoju metodu. Tam tikrais atvejais skiriamasis metodas yra diktuojamas pačios kiekybinių pažymių prigimties. Pav., kai kiekybinės dydžio grupės ar klasės sudaromos, tarsime, įmonėms pagal darbininkų skaičių, namams pagal butų arba kambarių skaičių, žemės ūkiams pagal gyvulių skaičių ar pn., tai kitoks metodas kaip skiriamasis nė neįmanomas. Nes vistiek kaip didelius intervalus čia paimsime, kiekviena sekanti kiekybinė grupė turės būti visu vienetu didesnė, kaip aukštutinę ribą apibrėžiantis skaičius žemesnėj grupėj. Pav., jei viena įmonių grupė sulig darbininkų kiekiu bus apibrėžta intervalu nuo vieno ligi, tarsime, 5 darbininkų, tai sekanti pagal įmonių didumą grupė, intervalo pagrindan padėjus taip pat 5 darbininkus, turės prasidėti 6 ir baigtis 10 darbininkų. Toliau eis grupė nuo 11 ligi 15 darbininkų, ir t. t. Tačiau, jeigu tos pačios įmonės būtų klasifikuojamos ne pagal dirbančių darbininkų kiekį, o pagal jų metinės gamybos vertę ar pagrindinio kapitalo dydį, tai čia skiriamasis metodas, jeigu intervalo skaitomuoju vienetu paimtume 5 tūkstančius litų, duotų šitokias grupes:

Nuo	1	ligi	5	tūkst.	lit.
„	6	„	10	„	„
„	11	„	15	„	„

Tačiau iš tokio dydžio grupių aprėžimo nebūtų aišku, ar įmonių, kurių kapitalas didesnis kaip 5000 litų, bet mažesnis kaip 6000 — visai nėra, ar yra, tik nežinia, kur tokios įmonės priskirtos: prie pirmosios dydžio grupės, ar jau prie sekančios. Išeities tokiu atveju ieškoma intervalo skiriamuosius mato vienetus sumažinant, pav., užuot 1000 lt. imama tik 100 litų (arba vieną litą, vieną centą). Intervalo skiriamuoju vienetu paėmus, pav., vieną litą, pirmoji dydžio grupė paliktų aprėžta lygiai taip pat kaip ir anksčiau (nuo 1000 ligi 5000), bet antroji prasidėtų ne 6000, o — 5001 litu, o trečioji — 10001 litu ir t.t. Bet ir šiuo atveju sumažėtų tik neaiškumo laipsnis ir galimos paklaidos didumas, bet ne klasifikacijos metodologiškai technikinė sunkenybė: nes ir čia liks vieno lito skirtumas, kuris, gali būti,

nulėmė mažesnio ar didesnio kiekio skaitomųjų vienetų priskyrimą prie vienos ar kitos kiekybinės grupės.

Bet tokiais atvejais, kai kiekybinės klasifikacijos pagrindan imamieji mato vienetai pareiškiami ir trupmenomis (pav., vertė arba kaina — litais, centais ir jų dalimis, plotai — hektarais, arais ir jų dalimis, amžius — metais, mėnesiais, dienomis, valandomis ir jų dalimis ir t. t.), neaiškumų palieka ir vadinamas jungiamasis metodas. Imkime tas pačias dydžio grupes litais ir aprėžkime: nuo 1000 ligi 5000 lt, nuo 5000 ligi 10000, nuo 10000 ligi 15000 ir t. t. Ir šitoki dydžio grupių aprėžimai be tolimesnių paaiškinimų dar neparodys, ar, sakysime, įmonė su 5000 litų savaitinės ar mėnesinės gamybos buvo priskirta prie pirmosios (mažiausios) dydžio grupės, ar pateko į antrąją. Dėl šios priežasties nei oficialiniuose statistikos šaltiniuose, nei privatinių tyrinėtojų darbuose nėra pasiekta vienodumo kiekybinių grupių intervalus aprėžiant. Šį klausimą suvienodinti yra pasiėmęs Tarpt. Statistikos Institutas, kuris 1938 m. Prahos kongrese turėjo išklausyt specialinės komisijos raportą ir padaryt tam tikrus nutarimus. Bet kongresas dėl politinių įvykių darbus nutraukė. Aplamai imant, statistikos praktikoj jungiamasis intervalo aprėžimo metodas, jei nieko kito nėra paaiškinta, paprastai šitaip vaizduojamas:

Amžiaus grupės skiriamuoju ir jungiamuoju metodais apibrėžtos.

Skiriamasis metodas	Jungiamasis metodas
0— 4	0 5
5— 9	5 10
10—14	10 15
ir t. t.	

Tokiais atvejais suprantama, kad žemutinė kiekvienos grupės intervalo riba apima visus tos grupės vienetus, kurie prasideda tokiu dydžiu, koks yra intervalo pradžioje, o aukštutinė intervalo riba apima visus vienetus, kurių dydis kiek galint prisiartina, bet dar nepasiekia, mato skaičių, kuris aukštutinę intervalo ribą aprėžia.

6. Statistikos lentelių sudarymo pavyzdžiai.

a) Paprastos esminėmis žymėmis sudarytos lentelės.

St. lentelės gali būti paprastos ir mažiau ar daugiau kombinuotos.

1) Paprastomis st. lentelėmis yra vadinamos tokios lentelės, kur skiltys tesudaro tik vieną visumą ir antgalvis sudarytas taip, kad kiekviena skiltis atžymima tik vienu pažymiu, o vienos eilutės atskirose skiltyse įrašyti duomenys, juos susumavus, atitinka visumą. Tokios lentelės pavyzdžiu gali būti kad ir lyties ar tautybės pažymiais sudarytoji lentelė, kur skilčių turės būti tiek, kiek yra atskirų pažymių, plius viena skiltis visumai, o eilučių tiek, kiek turės tas vienas pažymys atskirų dalinių (žiūr. lenteles 1 ir 1a).

Lentelė 1.

NATŪRINIS GYVENTOJŲ JUDĖJIMAS 1930—1936 MT.

Metai	Jungtuvės	Gimimai	Mirimai	Natūrinis prieaug.
1930	18 299	64 164	37 151	27 013
1931	18 840	63 419	37 478	25 941
1932	18 814	65 371	36 577	28 794
1933	19 511	62 145	32 749	29 396
1934	18 246	60 770	35 789	24 981
1935	18 254	57 970	34 595	23 375
1936	18 843	60 446	33 440	27 006

Lentelė 1a.

GIMUSIEJI 1936 MT.

	Berniukai	Mergaitės	Iš viso
Kaimuose	26 280	24 842	51 122
Miestuose	4 819	4 505	9 324
Iš viso	31 099	29 347	60 446

2) Kombinuota st. lentelė antgalvyje arba prieskiltyje privalo turėti vietos įrašams daugiau negu vienos žymės pažymiams. Pav., gali būti tokių lentelių, kur antgalvyje arba prieskiltyje iš karto yra kombinuojama vienos pagrindinės žymės pažymiai su kurios nors kitos žymės pažymiais. Pav., tautybės pažymiai antgalvyje kombinuojami su lyties arba amžiaus pažymiais, o prieskiltyje ploto pažymiai su laiko ar kitokiais pažymiais (žiūr. lentelę 2).

Lentelė 2.

GIMUSIEJI 1934—1936 MT.

		Moterystėje		Nemoterystėje		Iš viso		
		Berniukai	Mergaitės	Berniukai	Mergaitės	Berniukai	Mergaitės	Viso
1934:	a) kaimuose . .	24 795	23 432	1 706	1 633	26 501	25 065	51 566
	b) miestuose . .	4 242	3 987	504	471	4 746	4 458	9 204
	Iš viso	29 037	27 419	2 210	2 104	31 247	29 523	60 770
1935:	a) kaimuose . .	23 616	22 458	1 650	1 584	25 266	24 042	49 308
	b) miestuose . .	4 104	3 710	451	397	4 555	4 107	8 662
	Iš viso	27 720	26 168	2 101	1 981	29 821	28 149	57 970
1936:	a) kaimuose . .	24 632	23 226	1 648	1 616	26 280	24 842	51 122
	b) miestuose . .	4 367	4 064	452	441	4 819	4 505	9 324
	Iš viso	28 999	27 290	2 100	2 057	31 099	29 347	60 446

b) Kiekybiniais pažymiais sudarytos st. lentelės.

1. Statistiškai matematiniu metodu paruošiamos klasifikacijos dydžio grupių lentelės 1 ha intervalais galima rasti leidinyje Visuotinis Lietuvos Žemės Ūkio Surašymas 1930 m. I tomo 28 psl., kur yra patiekti suvestiniai duomenys Mažeikių, Šakių ir Ukmergės apskritims. Sekančiame (29) psl. sulig tos lentelės duomenimis yra nubrėžtos

kreivės, o p. p. 24—27 taip pat ir kitose to paties leidinio lentelėse yra sudarytos dydžio klasės jau su ne vienodais intervalais, bet kiek galint atsižvelgiant charakteringesniųjų paruošiamosios dydžio grupių lentelės bei kreivių rodomų susikaupimų (dispersijos). Geresnių pavyzdžių statistiškai matematiniu metodu vykdomai kiekybinei klasifikacijai ir nebereikalinga. Norėdami tuo pačiu metodu sudaryti mažesnį grupių skaičių, intervalus turime didinti. Pav., iš tos pačios lentelės įsitikinę, kad iš 38 696 ūkių bendrojo skaičiaus nėra smulkesnio ūkio kaip 1 ha ir stambesnio kaip 793,32 ha, norėdami sudaryti dydžio grupes su 5 ha intervalu kiekvieną, gautume 159 grupes, iš kurių paskutinioji apimtų ūkius su 790—793,32 (tikrenybėj, šioj grupėj tebūtų tik 1 ūkis su 793,32 ha). Jeigu norėtume grupių skaičių sumažinti ligi 100, kiekvienos grupės intervalui tektų paimti 8 ha (griežtai — 7,9332 ha). O norėdami tuos pačius duomenis suvesti į lentelę iš 10 dydžio grupių, intervalui turėtume imti jau 80 ha ir tuo būdu gautume šitokią lentelės schemą:

1)	Ūkiai	nuo	1— 80	ha
2)	„	„	80—160	„
3)	„	„	160—240	„
4)	„	„	240—320	„
5)	„	„	320—400	„
6)	„	„	400—480	„
7)	„	„	480—560	„
8)	„	„	560—640	„
9)	„	„	640—720	„
10)	„	„	720—800	„

Nesunku įsitikinti, kad tokiais dideliais intervalais sudarius dydžio grupes į pirmąją sueitų daugiau 99% visų ūkių ir daugiau kaip 90% visų žemės plotų. Tokios dydžio grupės būtų jau nepavaizdžios. Čia ir pasirodo reikalas dydžio grupes statistiškai matematiniu metodu sudaryti nedideliais intervalais, o didesniam pavaizdumui iš grupių jau sudaryti dydžio klases su nevienodais intervalais, atitinkančiais dispersijos pobūdį.

2. Galtono metodu iš tų pačių duomenų sudarant dydžio grupes ir apsistojus prie 10 grupių, kiekvienai grupei tektų paimti po 3.870 ūkius, nors paskutinei (dešimtajai) paliktų tik 3 866 ūkiai, t. y. 4 ūkiais mažiau, nes kai ūkių yra 38696, tai dalmuo (38696 : 10 = 3869,6) gaunamas su trupmena. O skaldyti ūkių negalima. Šio metodo pritaikymas parodyta patiekiamoj lentelėj 3.

Lentelė 3.

MAŽEIKIŲ, ŠAKIŲ IR UKMERGĖS ŽEMĖS ŪKIAI SUDOROTI Į DYDŽIO GRUPES GALTON'O METODU:

Nr.	Dydžio grupės ha	Ūkių skaičius	Ūkių plotas ha	Ploto % %
0	1	2	3	4
1	1— 3,2	3 870	7 928,46	1,47
2	3,2— 5,0	3 870	15 190,81	2,82
3	5,0— 6,5	3 870	22 824,62	4,23
4	6,5— 8,1	3 870	28 203,70	5,23
5	8,1— 9,7	3 870	34 192,83	6,34
6	9,7—11,7	3 870	41 969,50	7,79
7	11,7—15,5	3 870	49 773,25	9,23
8	15,5—19,4	3 870	62 835,63	11,68
9	19,4—28,9	3 870	91 509,07	16,98
10	28,9 ir daugiau (793,4)	3 866	184 635,24	34,23
	Iš viso	38 696	539 063,11	100,00

Lentelė 4.

KUMULIATYVINIŲ DYDŽIO GRUPIŲ SUDARYMO SCHEMA.

(Mažeikių, Šakių ir Ukmergės žemės ūkiai iš 1930 m. surašymo).

Nr.	Dydžio grupės ha ir jų kumuliavimas	Ūkių skaičius	Ūkių sk. % %	Ūkių plotas ha	Ploto % %
0	1	2	3	4	5
1	**1—2 ha**	**1 813**	**4,68**	**2 478,12**	**0,46**
2	2—3 „	1 612	4,17	3 854,95	0,72
2a	**1—3 „**	**3 425**	**8,85**	**6 333,07**	**1,18**
3	3—4 „	2 298	5,94	7 876,96	1,46
3a	**1—4 „**	**5 723**	**14,79**	**14 210,03**	**2,64**
4	4—5 „	2 011	5,20	8 889,24	1,65
4a	**1—5 „**	**7 734**	**19,99**	**23 099,27**	**4,29**
5	5—6 „	2 471	6,39	13 419,63	2,49
5a	**1—6 „**	**10 205**	**26,38**	**36 518,90**	**6,78**
6	6—7 „	2 919	7,54	1 880,97	3,49
6a	**1—7 „**	**13 124**	**33,92**	**38 399,87**	**10,27**
7	7—8 „	2 230	5,76	16 564,95	3,07
7a	**1—8 „**	**15 354**	**39,68**	**54 964,82**	**13,34**
8	8—9 „				
8a	**1—9 „**				
	Ir t. t.				

c) Klaidingos st. klasifikacijos st. lentelių pavyzdžiai.

Pagrindinis statistinės klasifikacijos reikalavimas, kad statistikos lentelėje patiektųjų statistinės masės grupių visuma sudarytų statistinį, t. y. vienos žymės visus pažymius apimantį vienetą. Šitam dėsniui nusižengta 1923 mt. visuotinio gyventojų surašymo duomenis klasifikuojant tautybės pažymiais.

Lentelė 5.

1) Klaidinanti lentelė apie Lietuvos gyventojus tautybėmis.

(Paimta iš lentelės „Lietuvos gyventojai“ ps. 37).

	Gyventojų skaičius	% %
Lietuviai	1 701 863	83,9
Žydai	153 743	7,6
Lenkai	65 599	3,2
Rusai	50 460	2,5
Vokiečiai	29 231	1,4
Latviai	14 883	0,7
Baltgudžiai	4 421	0,2
Kiti	1 592	0,1
Svetimšaliai	7 179	0,4
Iš viso	2.028 971	100,0

Šios lentelės pagrindinė klaida yra ta, kad joje dviejų žymių pažymiai yra suplakti į vieną: tautybės pažymiais yra sudarytos grupės tik Lietuvos piliečiams, o svetimšalių grupė tautybės pažymiais visai neišskirstyta. Be to, statistinėj lentelėj ši grupė nuo tautybinių grupių neatskirta ir užima tokią pat lentelėje eilutę kaip ir bet kurios tautybės grupė.

2) Klaidinga klasifikacija dėl skirtingos prigimties pažymių suplakimo.

Patiektoji lentelė turi parodyti įvairių rūšių susisiekimo priemonių skaičių ir sudėtį atskirais metais. Tačiau įvairias susisiekimo priemonių rūšis apibrėžiančių eilučių tarpe yra viena eilutė, kuri rodo ne priemonių skaičių, o susisiekimo priemones varančių jėgų rūšį (privatinių savininkų arkliai). Be to, pradedant 1932 mt., dvi susisiekimo priemonių rūšys (sunkieji vežimai ir resoriniai vežimai) yra sudėti į vieną grupę su arkliais, ir tuo būdu dviejų rūšių susisiekimo priemonės (vežimai) ir arkliai kartu rodomi susisiekimo priemonėmis.

Lentelė 6.

SUSISIEKIMO PRIEMONĖS KAUNE 1919—1934 M.).

(Paimta iš „Kauno miesto statistikos metraščio 1934 mt." psl. 25).

Priemonių pavadinimas	1929	1930	1931	1932	1933	1934
Lengvų vežimų	115	112	114	108	115	122
Sunk. vežimų	407	243	127			
Priv. sav. arklių	393	583	947	743	881	748
Resorinių vežimų	27	70	72			
Dviračių	3583	3708	3645	3184	3059	2795
Automobilių lengvų	263	215	388	286	299	326
Automobilių sunkių	61	91	109	99	81	84
Autobusų	44	56	56	58	52	55
Motociklų	131	93	200	182	166	179
Taksi	106	118	—	104	96	84

3) Klaidinga klasifikacija dėl skilčių turinio netikslaus apibrėžimo ir dėl sudėtinių dalių nevienodos prigimties.

Lentelė 7.

ŽALIAKALNIO KELTUVO KELEIVIŲ SKAIČIUS 1934 MT.

(Paimta iš „Kauno miesto statistikos metraščio 1934 mt." psl. 69).

Mėnesiai	Keleivių skaičius suaugusių	vaikų ir mok.	Viso
Sausis	70 872	14 289	85 161
Vasaris	67 371	13 918	81 289
Kovas	75 049	13 119	88 168
Balandis	73 316	10 115	83 431
Gegužis	73 085	15 291	88 376
Birželis	68 654	11 347	80 001
Liepos	43 824	8 728	52 552
Rugpiūtis	66 997	10 919	77 916
Rugsėjis	75 627	16 478	92 105
Spalių	57 271	13 217	70 488
Lapkritys	72 396	25 500	97 896
Gruodis	68 681	17 762	86 443
Abonentiniai bil.	500	3 500	4 000
Viso	813 643	174 183	987 826

Šioj lentelėj skilčių turinys apibrėžtas neaiškiomis sąvokomis, nes amžiaus grupės (suaugę, vaikai) privalo būti apibrėžiamos objektyviniu pažymiu, t. y. metais. Be to, vaikai suplakti į vieną grupę su mokiniais, o mokinių tarpe yra ir suaugusių žmonių (studentai).

Taip pat statistinės lentelės eilučių pažymiai dvejopu principu sudaryti: skaitomaisiais laiko periodais (mėnesiais) ir viena eilutė dar atskirai. Be to, šios eilutės turinys iš esmės skiriasi nuo visų kitų eilučių turinio (ten keleiviai — čia abonentiniai bilietai, su kuriais keleivis važinėja daugiau ne kaip vieną kartą). Dėl to ir visų eilučių skaičius sumuoti, kaip vienareikšmius, nėra galima.

4) Klaidinga st. lentelė dėl pažymių prigimties nevienodumo ir sąvokų nelogiškumo.

Lentelė 8.

MIESTO BUTŲ KOLONIJŲ GYVENTOJAI 1934 METAIS.

(Paimta iš „Kauno miesto statistikos metraščio 1934 mt." psl. 30).

Butų kolonijos	Vyrų	Moterų	Vaikų	Viso asmenų	Šeimų
I	60	75	138	273	61
II	39	126	129	294	67
III	2	2		4	2
IV	34	38	44	116	26
V	12	27	26	65	15
Viso	147	268	337	752	171

Šioj lentelėj nėra klasifikacijos vienodumo: viena skaitomųjų vienetų dalis suskirstyta lyties pažymiais, kita tos pačios kilmės masės dalis (vaikai) lyties pažymiais neklasifikuota. Be to, „vaiko" sąvoka jokiu objektyviniu pažymiu neapibrėžta ir nežinia, ar tai yra tik kūdikiai, ar turintįs nedaugiau kaip 10 mt., ar gal 15 ar 16 mt. Taip pat paskutinėj skiltyj parodytieji šeimų skaičiai iš lentelės antgalvio įrašų palieka nesuprantami: nes nėra aišku, ar prieš tai buvusioj skilty

bendrasis visų asmenų skaičius tose butų kolonijose liečia tik pavienius asmenis, jokios šeimos nepriklausančius, ar gal tos šeimos apima visus pirmesnėse skiltyse parodytus asmenis.

7. Statistikos duomenų vaizdavimas grafikos priemonėmis.

Vienos ar kitos rūšies pažymiais suklasifikuoti statistikos duomenys bet kurios formos statistikos lentelėse kad ir įgalina statistinės masės charakteringas ypatybes pažinti, tačiau lentelėse patiektieji duomenys yra reikalingi gero į juos įsižiūrėjimo arba stebėjimo. O gali būti reikalas tik pačius bendruosius statistinės masės bruožus pažinti, ir tai kiek galint greičiau. Vadinasi, statistikos lentelėse patiektųjų duomenų nuodugniam studijavimui gali nebūti laiko, pagaliau, nei reikalo.

Kiekvienos statistinės masės pačius bendruosius bruožus visai ryškiai parodo mokamai paruoštas grafiškas vaizdavimas. Grafiškai vaizduojami duodasi bet kurios rūšies statistikos duomenys. Tik kad grafika iš tikrųjų pačius charakteringus bruožus iškeltų, yra reikalinga grafiškojo vaizdavimo formą ir priemones kiekvienos statistinės masės ypatybėms pritaikinti.

Jau anksčiau mes turėjome progos pažinti, kad sulig tuo, ar statistinę masę sudaro kurie nors subjektai — daiktai, dalykai, ar tik reiškiniai — įvykiai, atsitikimai, — yra reikalinga ir statistinio stebėjimo metodus pritaikinti. Tą patį tenka pasakyti ir dėl statistinio vaizdavimo priemonių, ypač dėl grafiškojo vaizdavimo priemonių. Jei grafiškai norima atvaizduoti kurios nors statistinės masės stovis, sudėtis ar atskirų jos dalinių padėtis bei savitarpio lyginamasis svoris, — reikalinga ir atitinkamos grafiško vaizdavimo formos parinkti. Ir kitokia grafiško vaizdavimo forma bus reikalinga parinkti, jei norėsime atvaizduoti kurio nors masinio reiškinio kitimo charakteringas ypatybes laiko bėgyje.

Stovio arba padėties statistikai (masių stoviui ir sudėčiai) grafiškai atvaizduoti geriausia tinka plotiniai vaizdai: diagramos, kvadragramos ar kitoki geometrijos taisyklių priklausantys vaizdai. Lengviausia plotinio vaizdavimo forma, — o paprastų statistikos lentelių duomenims vaizduoti ir pavaizdžiausia, — tai taisyklingų keturkampių ar pailgų, bet tiesiais kampais, plotų vaizdai.

Tokie kvadratinio matavimo dėsnių priklausantys vaizdai pirmiausia prašyte prašosi visus tuos statistikos duomenis atstovauti, kurie turi reikšti atskirų žemės plotų dydį ir jų savitarpio santykį (pasėlių plotai, žemės ūkių atskirų dydžių grupių lyginamasis svoris ir t. t.). Bet šios rūšies plotiniai vaizdai tinka ir gyventojų stovio statistikos duomenims, taip pat visur ten, kur tik nėra reikalo kitimus laiko bėgyje atvaizduoti. Tačiau plotiniai vaizdai tinka atvaizduoti ir bet kurios statistinės masės stovio arba sudėties pasikeitimus tam tikram laikui praėjus, pav., atskirų surašymų duomenims palyginti.

Aplamai, keičiant piešinį arba spalvas, paprasti kvadratinio matavimo plotiniai vaizdai duoda labai daug įvairių vaizdavimo galimybių bei formų. Dar daugiau tokių plotinio matavimo formų patiekia komplikuotesni vaizdai arba vadinamos geometrinės figūros, pradedant visai taisyklingomis diagramomis (skrituliais, trikampiais ir t. t.), baigiant įvairiausių formų vaizdais, kurie grafiško vaizdavimo praktikoj dažnai įgauna formą miniatiūrinių, vaizduojamos masės rūšį atitinkančių, atvaizdų. Pav., gyventojų statistikos duomenys suskaidyti bet kurios rūšies pažymiais, vaizduojami miniatiūrinėmis įvairaus didumo arba lyginamąjį svorį atitinkančio skaičiaus žmogaus figūromis. Taip pat gyvulių statistikos duomenys vaizduojami atitinkamos gyvulių rūšies paveikslais, trobesių statistika — trobesiais, kurių nors produktų ar prekių statistikos duomenys — produktais ar prekėmis, ir t. t. Jei norima grafiškai atvaizduoti bet kurios masės pasiskaidymą atskiruose geografiniuose rajonuose, patiekiamos vadinamos kartogramos, kur ir vėl skirtingos piešinių formos arba skirtingos spalvos bei įvairiausios figūros įgalina perduoti ir atvaizduoti net ir labai sudėtingų statistinių masių ypatybes.

Tačiau vis viena kokios rūšies plotiniai vaizdai būtų naudojami, visais atvejais tenka tų vaizdų formas pagrįsti geometrinio skaičiavimo dėsniais, t. y. kiekvieną plotinį vaizdą ar to vaizdo bet kurią dalį didinant ar mažinant, pasikliauti geometrinės proporcijos dėsniu.

Judėjimo statistikos duomenims, t. y. masinių reiškinių arba atsitikimų kitimams grafiškai atvaizduoti, tinka-

miausia forma yra vadinamos liniagramos, t. y. kreivės. Kreivių piešimo technika taip pat duoda gana daug įvairių galimybių masinių reiškinių kitimams vaizduoti.

Tam tikros rūšies reiškiniams grafiškai atvaizduoti gali būti liniagraminis vaizdavimas kombinuojamas su plotiniu vaizdavimu. Pav., gimimų ir mirimų kitimus vaizduojant paralelinėmis kreivėmis, šios dvi kreivės niekuomet nesusikirs, — jeigu visą laiką mirimų bus mažiau kaip gimimų. Tuo būdu tarp abiejų kreivių likęs plotas automatiškai atvaizduoja natūrinio gyventojų prieauglio masę. Jei gi mirimų ir gimimų kreivės susikirstų, t. y. kuriais nors metais mirimų būtų daugiau kaip gimimų, tai atitinkamas plotas, kitokios rūšies braižiniu pavaizduotas, parodys natūrinio gyventojų sumažėjimo masę. Panašiai galime atvaizduoti kombinuojant kreivę su plotiniais vaizdais ir užsienių prekybos balanso kitimus, biudžeto pajamas ir išlaidas kartu su biudžetiniais pertekliais ir deficitais, ir t. t.

Kad grafiškas vaizdavimas būtų statistiškai prasmingas ir savo tikslo pasiektų, yra reikalinga prisilaikyti šių taisyklių:

1) Kiekvienas grafikos vaizdas turi būti ryškus ir visa ką turi parodyti patsai piešinys.

2) Visi reikalingieji skaičiai privalo būti čia pat piešinyje arba gretimai įrašyti. Taip pat ir reikalingieji atskirų piešinio dalių sutartiniai pažymiai ryškiai ir matomoj vietoj turi būti atžymėti.

3) Pavaizdžiausi ir ryškiausi yra paprasčiausieji plotiniai vaizdai. Dėl to ir rekomenduojama vengti per daug komplikuoto figūrinio vaizdavimo, kur figūrinių vaizdų netaisyklingumas savaime mažina perspektyvą arba pavaizdumą ir apsunkina vaizduojamos masės arba reiškinių atskirų dalinių lyginamąjį svorį teisingai įvertinti.

4) Grafiškas vaizdavimas turi atitikti ir estetikos reikalavimus. Vadinasi, bet kurios rūšies plotiniai vaizdai arba liniagramos savaime, taip pat jų deriniai (kombinacijos), turi estetiškai atrodyti.

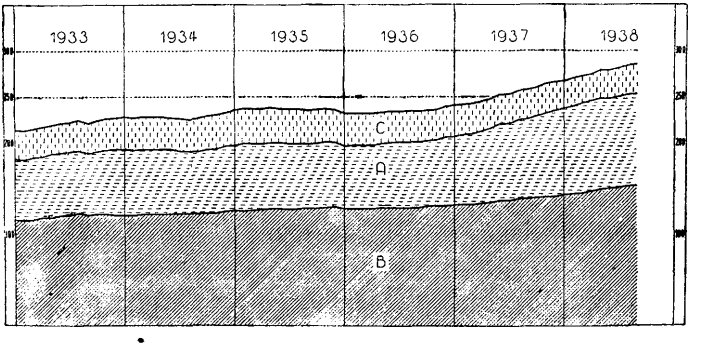
5) Grafiškųjų vaizdų statistinį prasmingumą ir jų estetinę formą daug nulemia tikslingas ir statistiškai prasmingas m a s t e l i o parinkimas. Sulig grafiškai vaizduojamų duomenų ypatybėmis, tam tikrais atvejais gali būti prasminga panaudoti ir vadinamas l o g a r i t m i n i s m a s t e l i s.

6) Logaritminis mastelis statistiškai prasminga yra naudoti tais atvejais, kai grafiškai vaizduojamų statistinės masės dalinių skirtumai yra labai dideli, ir dėl to perdėm išlaikytas paprastasis mastelis labai daug vietos užimtų ar duotų estetiškai negražų vaizdą. Be to, logaritminis mastelis statistiškai prasmingas gali būti tais atvejais, kai vaizduojamos statistinės masės kokybinių arba kiekybinių dalinių skirtumai arba kitimai laiko bėgyje vyksta geometrinės progresijos ar geometrinės proporcijos dydžiais.

Patiekiamieji grafiško vaizdavimo pavyzdžiai apibrėžtąsias tezes ryškiau paaiškins.

Diskontas ir paskolos (mil. Lt)

A—7 Kauno bankai
B—Žemės Bankas
C—Taupomosios Valstybės Kasos



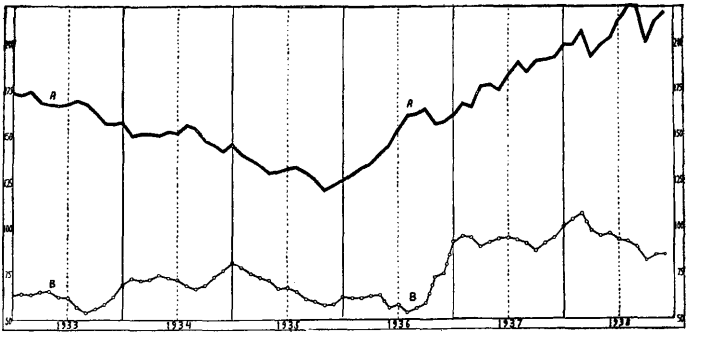
(Klišė iš Lietuvos Banko biuletenio)

Čia patiekiama kombinuota ploto ir linijų diagrama. Kreivės vaizduoja kitimų tendenciją (diskontą ir paskolas) atskirose kredito įstaigose. O tarpai tarp kreivių, skirtingai kiekvienas nupaišyti, parodo diskonto ir paskolų dydį, o tuo pačiu ir lyginamąjį svorį tose pačiose kredito įstaigose.

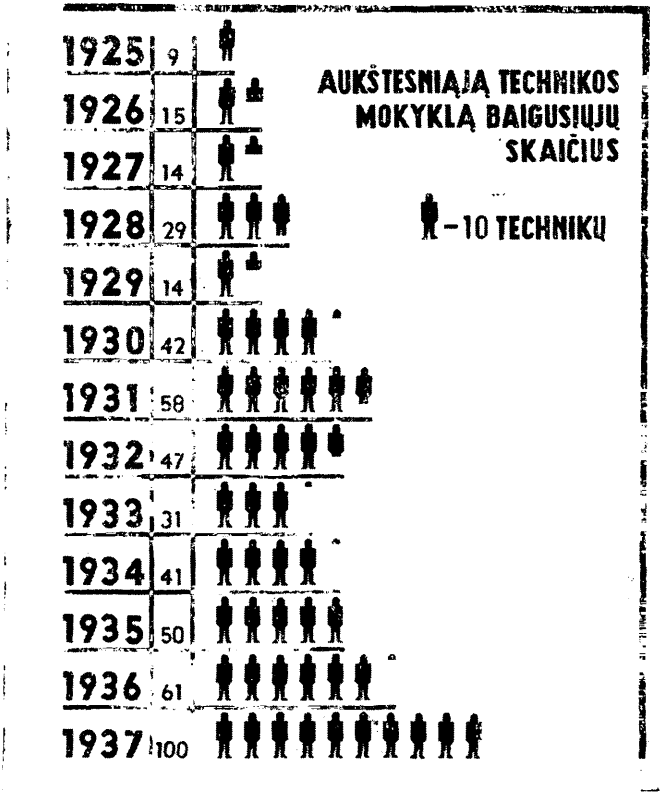
Ši diagrama yra pavaizdesnė, negu žemiau dedamoji, kur privatiniai ir valstybiniai indėliai atvaizduoti vien kreivėmis.

Privatiniai ir valstybiniai indėliai Lietuvos kredito įstaigose — be smulkaus kredito draugijų (mil. Lt)

A—Privatiniai B—Valstybiniai

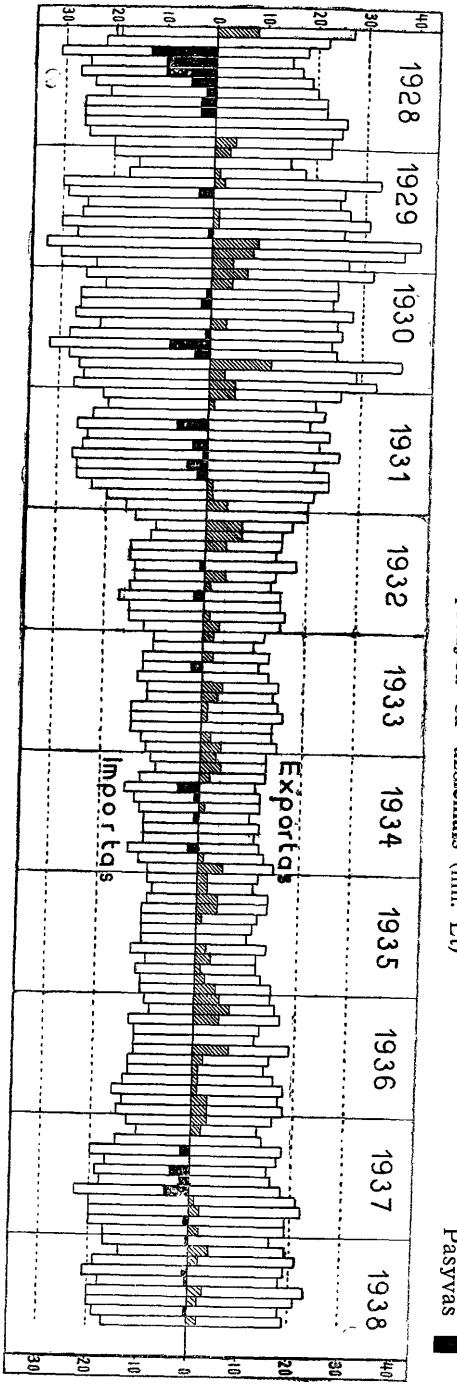


(Klišė iš Lietuvos Banko biuletenio)



(Klišė iš Susisiek. M-jos Metraščio)

Tikslingos figūrinės diagramos pavyzdys: kiekviena figūra atstovauja griežtai apibrėžtą statistinės masės dalinį.



(Klišė iš Lietuvos Banko biuletenio)

Čia patiekiama kombinuota plotinė diagrama, kur atvaizduota užsienių prekybos apyvarta atskirais mėnesiais, paėmus vaizdavimo baze tos prekybos balansą. Nuo bazės į abi puses einantieji stulpeliai rodo eksporto ir importo savitarpio santykį bei apimtį (dydį), o tamsesniais arba visai juodais plotiniais bruožais tuo pačiu kartu parodomas ir to santykio rezultatas (eksportas didesnis už importą, ar atvirkščiai).

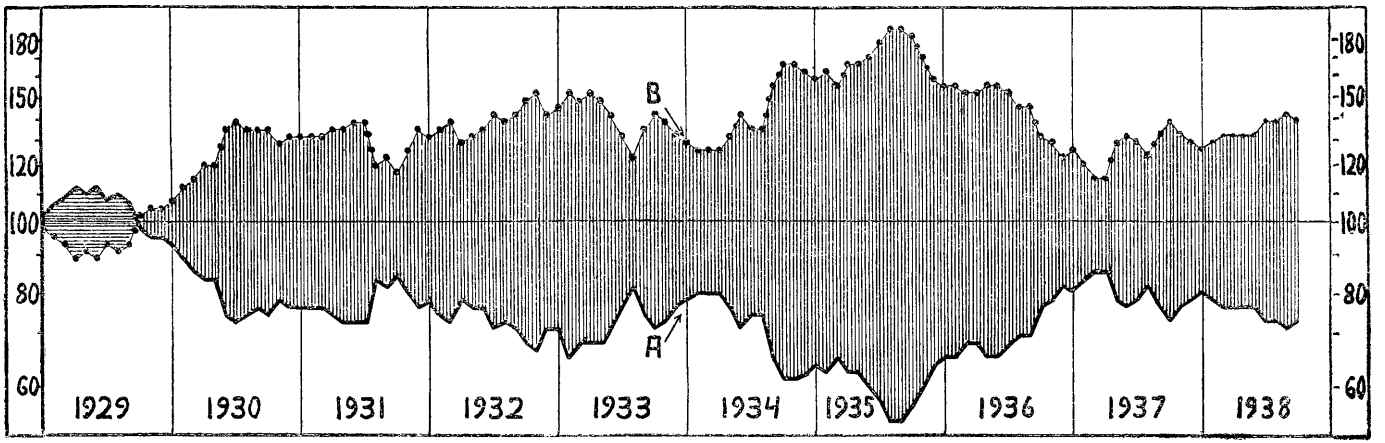
129 pusl. diagramoj vad. kainų žirklės, atvaizduotos logaritminiu masteliu. Žemės ūkio produktų perkamoji galia (A) apibrėžiama dalmeniu, kuris gaunamas žemės ūkio produktų kolektyvinį in-

Žemės ūkio ir pramonės savitarpio perkamoji galia.

(1926—1929=100)

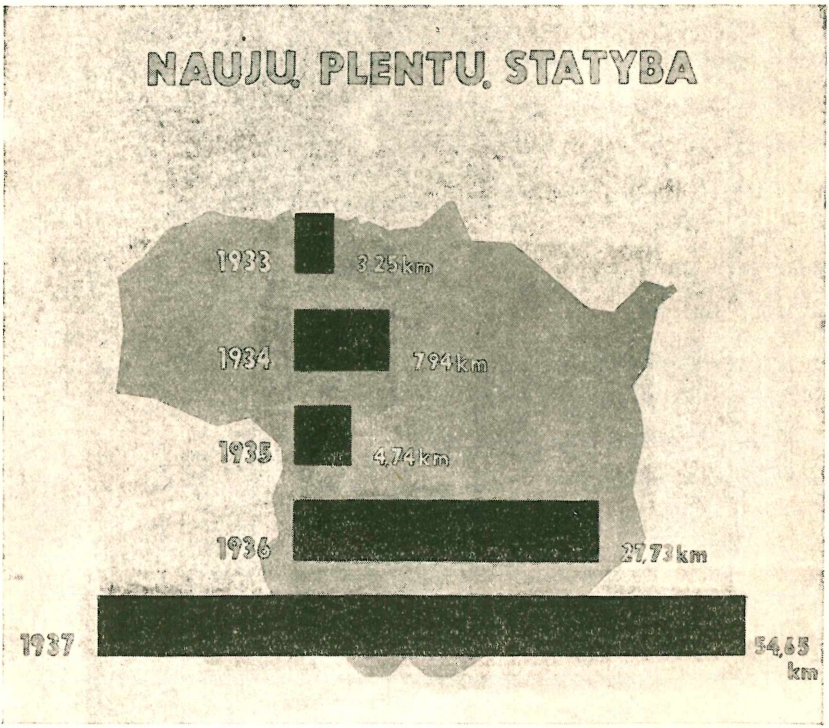
(Logaritminis mastelis)

A—Žemės ūkio perkamoji galia **B**—Pramonės perkamoji galia



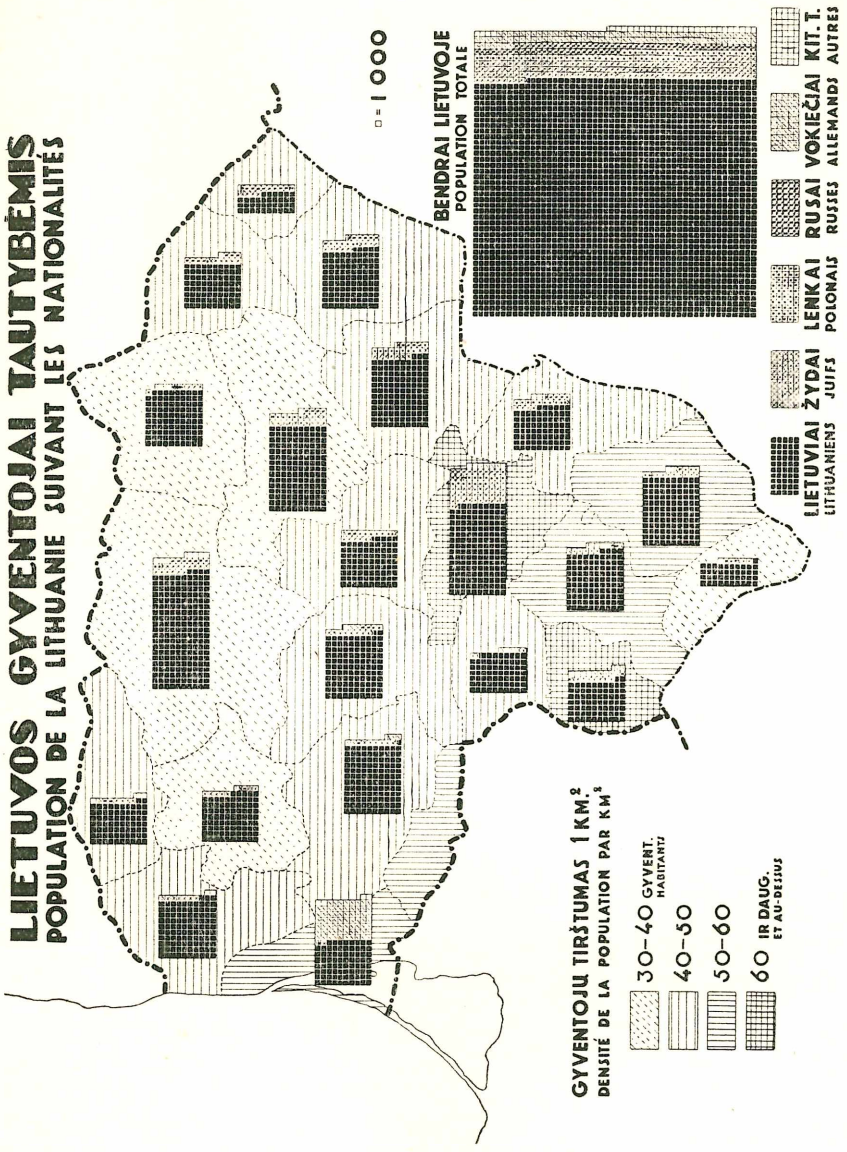
(Klišė iš Lietuvos Banko biuletenio)

deksą dalinant pramonės dirbinių kolektyviniu indeksu ir atvirkščiai: pramonės dirbinių perkamoji galia apibrėžiama dalmeniu, gaunamu pramonės dirbinių kolektyvinį indeksą padalinus žemės ūkio kainos kolektyviniu indeksu. Visais tais atvejais, kai abudu indeksai sutampa, žirklės visiškai susiglaudžia (dalmuo gaunamas 1 arba 100, jei bazė prilyginama 100). Logaritminis mastelis įgalina abidvi kreives tolintis ir artintis visai simetriškai. Jei būtų panaudotas paprastasis mastelis, tai viršutinė kreivė nuo bazinės linijos visuomet nutoltų daugiau kaip apatinė.



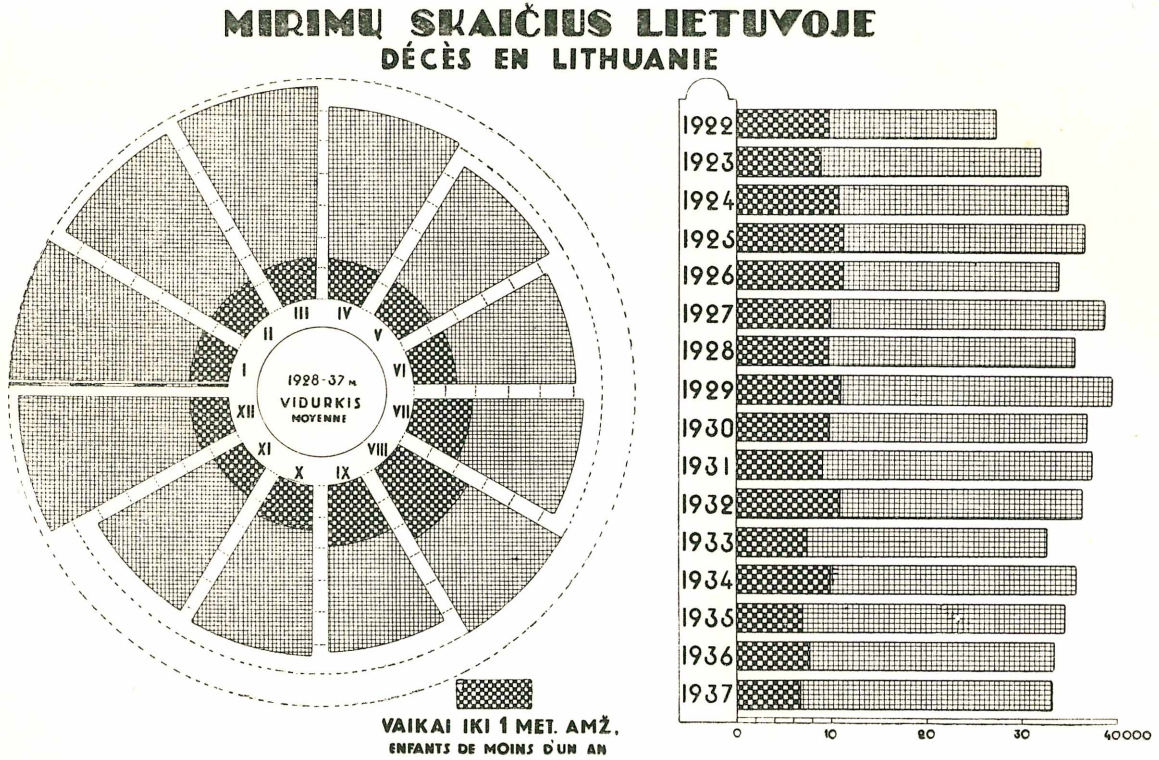
(Iš Sus. M-jos 1937 mt. metraščio).

K l a i d i n a n t i p l o t i n ė k a r t o g r a m a: naujų plentų statybai vaizduoti paimtas ploto mastelis nesuderintas su žemėlapio masteliu (sudėjus 1933—37 mt. naujai pastatytų plentų plotą gaunamas vaizdas, kad per tuos kelis metus plentais paversta apie ½ viso Lietuvos ploto).



(Klišė iš CSB Statistikos Metraščio).

Tikslinga plotinė kartograma: kiekvienoj apskrityj gyventojų tirštumas parodytas proporcionaliai su gyv. skaičiumi ir plotu.



(Klišė iš CSB Statistikos Metraščio).

Tikslingai sudaryta diagrama, kur parodyta reiškinių (mirimų) proporcija ir judėjimas laiko bėgyje.

KETVIRTOJI DALIS.

A. Statistinių eilučių santykis ir statistikos analizio metodai.

1. Statistinio analizio formos ir priemonės.

Tam tikra prasme statistinis analizis prasideda jau tuo metu, kai statistinio stebėjimo keliu gautieji duomenys imami klasifikuoti. Kokybiniais arba kiekybiniais pažymiais sugrupuotų statistikos duomenų atvaizdavimas statistikos lentelėse, arba diagramose, jau rodo statistinės masės skirtingų dalinių lyginamąjį svorį, o iš dalies ir jų savitarpio saitus. O savo prigimtyje tai ir yra jau analizio padarinys. Tačiau iš kitos pusės, ši pirmoji statistikos duomenų klasifikacija tesudaro tik sąlygas statistiškai prasmingam duomenų analiziui. Ypač tai tenka pasakyti apie klasifikaciją kokybiniais pažymiais. Kokybinės statistinės masės vienetų pažymiai, kurie parodo stebimo objekto vienetų rūšį, kilmę ar kitas jo ypatybes, ir būna stebimi, t. y. statistinio stebėjimo metu surašinėjami, kad, tais pažymiais statistikos masę suskirsčius į grupes, būtų galima pačias charakteringas ypatybes pažinti. Tik šios rūšies duomenis prieš akis turint jau galima pradėti ir tikrąjį statistinio analizio darbą, t. y. tyrinėjimą, kokiame griežčiau apibrėžtame santykyje atskiros žymės ir jų pažymiai visoje masėje yra pasiskaidę, kurios žymės ir pažymiai masėje vyrauja, o kas tik nereikšmingas išimtis tesudaro. Net ir kiekybinės klasifikacijos pirmojoj stadijoj, t. y. kai turimoji statistikos medžiaga kuriuo nors vienu griežtai išlaikytu principu suskaldoma į vienodų intervalų ar lygaus skaičiaus dydžio grupes, taip pat vertintina tik kaip paruošiamasis darbas ir priemonė statistiniam analiziui.

Pirmasis statistinio analizio uždavinys, — tai statistinės klasifikacijos priemonėmis patiektuosius duomenis paryškinti ir, kiek galint, prie vienodo vardiklio privesti, kad būtų galima

analizuojamos statistinės masės charakteringas ypatybes pastebėti ir jų savitarpio saitus įžvelgti. Toliau jau seka vadinamųjų statistikos dėsningumų apibrėžimas bei konstatavimas, atskirų statistinės masės dalinių, arba dviejų ar kelių statistinių masių, savitarpio ryšių bei pareinamybių ieškojimas ir statistiniais dydžiais, t. y. skaičiais, jų pareiškimas.

Pačių statistinio analizio priemonių yra gana daug — pradedant elementarinės aritmetikos veiksmais, baigiant algebros bei vad. aukštosios matematikos lygčių sistemomis, kurios, tačiau, statistikos analizio darbuose tik tiek naudotinos, kiek jos įgalina statistikos analizio darbą palengvinti bei sutrumpinti arba statistinių dydžių esmę bei statistinės eilutės prigimtį paryškinti.

Paprasčiausioji, bet kartu ir reikalingiausioji, statistinio analizio priemonė — tai statistikos eilučių statistinės lentelės (atskirų skilčių) santykį apibrėžiamieji ir matuojamieji skaičiai.

2. Statistikos eilučių santykį matuojamieji skaičiai.

Statistikos eilučių santykį savo esmėje vaizduoja bet kurios rūšies statistiniai skaičiai, kurie yra sudaryti statistinės klasifikacijos vienu ar kitu griežtai išlaikytu principu ir atvaizduoti statistikos lentelėje arba grafikos priemonėmis. Tačiau absoliutiniai statistikos skaičiai ne visuomet būna pakankamai pavaizdūs, kad vienu peržvelgimu būtų galima ne tik, tarsime, statistinių grupių lyginamąjį svorį matyti, bet ir to svorio laipsnį griežtai apibrėžti. Be to, pirmykštės klasifikacijos eigoje gautieji skaičiai ne visuomet yra vienareikšmiai, t. y. neturi bendro vardiklio. Šiuos du absoliutinių statistikos eilučių skaičių trūkumus turi tikslo pašalinti statistinis analizis s a n t y k i n i a i s s k a i č i a i s.

S t a t i s t i n i a i s a n t y k i o s k a i č i a i gali būti (ir būna) pareiškiami įvairiopais dydžiais: trupmenomis, procentais, promiliais, vadinamais indeksais, koeficientais ir t. t. Tačiau visus šiuos skirtingus ir visokiariopais vardais vadinamus santykio skaičius, kaip statistinio analizio priemonę, tenka skirti į dvi pagrindines grupes arba rūšis sulig tuo, ar statistinio santykio skaičiai vaizduoja tik v i e n o s statistinės masės s u d ė-

tines dalis, ar statistiniais santykio skaičiais norima atvaizduoti (arba vaizduojama) dviejų ar keletos statistinių masių savitarpio lyginamąjį santykį. Dėlto ir sutarsime tuos statistinius santykio skaičius, kurie vaizduoja vienos statistinės masės dalių savitarpio lyginamąjį svorį (procentais, promiliais ir kt.), vadinti dalmeniniais santykio skaičiais arba, sutrumpinus, tiesiog statistiniais daliniais, o tuos statistinio santykio skaičius, kuriais vaizduojama dviejų arba kelių statistinių masių (statistinių eilučių) pareinamybinį santykį — lygmeniniais santykio skaičiais arba tiesiog statistiniais lyginiais.

3. Statistiniai santykio daliniai.

Lentelė 9.

LIETUVOS GYVENTOJAI 1934 M. PRADŽIOJE TIKYBOMIS.

Tikybų pavadinimai.	Absoliutiniai skaičiai.	Statistiniai santykio daliniai °/₀°/₀
Romos katalikai	1 746 942	80,48
Evangelikai	207 118	9,54
Ortodoksai	55 122	2,54
Kiti krikščionys	1 889	0,09
Izraelitai	157 527	7,26
Kiti nekrikščionys	1 640	0,07
Nenurodytos tikybos	378	0,02
Viso	2 170 616	100,00

Statistinio santykio daliniais, arba dalmeniniais santykio skaičiais, gali būti vaizduojama kiekviena statistinė eilutė, kuri yra bet kurios ypatybės (žymės) pažymiais suskaidyta. Pav., gyventojų surašymo duomenys skaidomi tautybės pažymiais ir, tuo būdu, gaunamos lietuvių, žydų, lenkų ir kt. tautybių grupės. Kiekviena tautybė sudarys atskirą grupę, o šių grupių visuma parodys bendrąjį gyventojų skaičių. Tačiau, visi šie atskirų tautybių skaitomieji vienetai gali būti pareikšti ne tik absoliutiniais skaičiais, bet ir statistinio santykio daliniais, kurie kiekvieną tautybę atstovaus proporcionaliai. Ar šios rūšies proporciniai daliniai bus pareikšti trupmenomis, procentais, promiliais

ar kt., tai esmės nepakeičia. Svarbu, kad šie statistinės masės daliniai būtų kiek galint ryškesni, lengviau suvokiami ir ilgiau galvoj palaikomi.

Praktikoj statistiniai santykio daliniai dažniausiai pareiškiami procentais. Tokio statistinės masės dalinių procentais pareiškimo pavyzdžiu gali būti st. lentelė, kur 1923 m. Lietuvos gyventojų surašymo duomenys atvaizduoti tikybos pažymiais.

Patiektosios lentelės statistikos duomenys antrojoj skiltyj yra ryškesni ir lengviau prisimenami, o tikybų proporcija palieka ta pati. Vadinasi, bet kurios rūšies statistikos duomenys gali būti atvaizduoti ne tik absoliutiniais skaičiais, bet ir statistinio santykio daliniais. Tačiau iš to dar neseka, kad statistinio santykio daliniai gali absoliutiniais skaičiais sugrupuotus statistikos duomenis pakeisti. Jeigu mes, pav., paimsime dvejetą statistinių masių (dviejų valstybių gyventojų skaičius) ir suskaidysime tais pačiais pažymiais (tikybos, tautybės, verslo ar kt.), tai, prie vienodos abiejų masių statistikos dalinių proporcijos, pačius dalinius, pareiškus juos procentais ar kt., gausime tuos pačius, — nežiūrint į tai, kad viena iš vaizduojamų masių būtų dešimtį arba net šimtą kartų didesnė už antrąją. Arba imkime tą pačią statistinę masę iš dviejų laikotarpių, kai tos masės skaitomųjų vienetų bendrasis skaičius, arba kai kurių dalinių proporcija, žymiai pasikeitus. Tokių pakeitimų tikrosios prasmės nei kilmės nebūtų matyti, jei tą statistinę masę vaizduojantieji duomenys būtų pareikšti vien tik statistiniais santykio daliniais.

Iš to seka šis kiekvienu atveju įsidėmėtinas dėsnis: statistiniai santykio daliniai yra reikalingi lygiagrečiai patiekiamų absoliutinių skaičių visais tais atvejais, kai tie daliniai yra naudojami kaip statistinio analizio priemonė. Vadinasi, tais atvejais, kai statistiniais santykio daliniais tik pavaizduojama kuri nors šiaip jau žinomos ir pažįstamos statistinės masės ypatybė, galima tuos dalinius atitinkančių absoliutinių skaičių ir neminėti. O kai statistinio santykio daliniai lyginamajam metode tarnauja išrodomąja statistinio analizio priemone (tyrinėjamas statistinių masių arba atskirų tos pačios masės grupių lyginamasis svoris ar jų santykio kitimai lai-

ko bėgyje), — yra reikalinga patiekti ir absoliutinius skaičius. Priešingu atveju statistiniais santykio daliniais vaizdavimas gali būti ne tik neprasmingas ar netikslingas, bet ir klaidinantis.

Aplamai, jei sutarsime bet kurios statistinės masės vienetų bendrąjį skaičių žymėti raide X, o bet kurios žymės pažymiu (bet kurios tautybės, tikybos, amžiaus ar kt.) atžymėtose grupėse esančių vienetų skaičių raidėmis x_1, x_2, x_3,..., x_n, tai bet kurios x_i grupės bet kuri statistinio santykio dalinį d_i gausime pagal formulą:

$$d_i = \frac{x_i}{X} \qquad \text{(III)}$$

Iš to seka, kad:

$$\frac{x_1}{X} + \frac{x_2}{X} + \ldots + \frac{x_n}{X} = \frac{\Sigma x_i}{X} = 1.$$

O kadangi:

$$d_i = \frac{x_i}{X}, \quad \text{tai ir}$$

$$d_1 + d_2 + \ldots + d_n = \Sigma d_i = 1.$$

Jei statistinio santykio dalinius norėsime žymėti ne trupmenomis, o procentais, tai formulą teks pakeisti į:

$$d_i\,\% = \frac{x_i \cdot 100}{X} \qquad \text{(III a)}$$

Savaime suprantama, kad šiuo atveju ir statistinių dalinių suma

$$\Sigma d_i\,\% = 100.$$

4. Statistinio santykio lyginiai.

Statistiniai santykio lyginiai, priešingai statistinio santykio daliniams, visuomet rodo d v i e j ų statistinių masių, ar jų grupių santykį arba dviejų statistinių eilučių kurių nors dydžių arba pažymių santykį. Skirtinga čia ir statistinio skaičiavimo technika. Statistinio santykio daliniams paimtoji statistinės masės visuma prilyginama 1, 100, 1000 ar kt., o tos pačios masės atskiri daliniai pareiškiami trupmena, procentais, promiliais ar kt. Statistinio gi santykio lyginiams visuomet kuriuo nors kriterium sudaryta viena statistinė masė, ar jos dalinys, tarnauja baze kitai statistinei masei, arba jos daliniams, palyginti. O jeigu statistiniais santykio lyginiais norima atvaizduoti tos pačios statistinės masės kurį nors reiškinį arba pažymį, tai viena tos

pačios masės dalis imama lyginimo baze kitai tos pačios masės grupei.

Pavyzdžiu galime imti kad ir bet kurių metų gimimų duomenis. Štai 1936 m. Lietuvoje gimė 60 446 kūdikiai, iš jų buvo 31 099 berniukai ir 29 347 mergaitės. Gimusiųjų berniukų ir mergaičių atskirai paimtieji skaičiai sudarys visos gimusiųjų masės dalinius. Jei sutarsime šiuos statistinio santykio dalinius pareikšti procentais, tai berniukų statistinio santykio dalinį gausime $\frac{31\,099.100}{60\,446} = 51,45$, o mergaičių $= \frac{29\,347.100}{60\,446} = 48,55$. Tačiau tuos pačius berniukų ir mergaičių skaičius galima atvaizduoti ir statistinio santykio lyginiais. Pav., tą patį berniukų skaičių padauginę 100, galime padalinti bendruoju mergaičių skaičiumi. Rezultate gausime $\frac{31\,099.100}{29\,347} = 106,0$, t. y. lyginant su mergaitėmis, gimusieji berniukai sudaro 106%.

Lentelė 10.

GIMUSIŲJŲ SANTYKIO SKAIČIAI.

Metai	Gimimų santykio daliniai % %			Gimimų sant. lyginiai % %	
	Iš visų gimusių berniukų sudarė % %	Iš visų gimusių mergaičių sudarė % %	Iš viso	100 gimusių mergaičių teko berniukų	100 gimusių berniukų teko mergaičių
1932	51,49	48,51	100	106,2	94,2
1933	51,49	48,51	100	106,2	94,2
1934	51,42	48,58	100	105,8	94,5
1935	51,44	48,56	100	105,9	94,4
1936	51,45	48,55	100	106,0	94,4

Šiuo atveju vienas statistinės masės dalinys padaromas lyginimo baze kitam, t. y. kitu pažymiu atžymėtam statistinės masės daliniui. Tuo būdu išeina, kad gimusiųjų berniukų ir mergaičių skaičiai, statistinio santykio lyginiams skaičiuot sudaryti, gauna jau dviejų savarankiškų statistinių masių pobūdį. Ir šios statistinės masės lyginamos lygiai taip pat, kaip ir, pav., yra lyginami bendrieji gyventojų skaičiai su gyvenamu plotu, mo-

kinių arba mokytojų skaičiai su mokyklų skaičiumi, gimimai, mirimai ar kt. su bendruoju gyventojų skaičiumi ar kuriais nors jų daliniais ir t. t.

Aplamai, statistinio santykio lyginiai yra gana tiksli statistinio analizio priemonė. O kadangi statistiniais santykio lyginiais būna vaizduojama ne tik tos pačios kilmės statistinių masių daliniai, bet ir visai skirtingos prigimties statistiniai duomenys, tai statistinis analizis santykio lyginiais pasidaro itin svarbus ir atsakomingas. Pav., imkime statistinio santykio lyginiais vaizduojamą gyventojų plotą su gyventojų kiekiu, kai norima parodyti — tirštai ar skystai kurie nors plotai žmonių gyvenami. Kad šios rūšies lyginiai savo tikslą pasiektų, yra reikalinga, kad abidvi statistinės eilutės (gyventojų ir ploto duomenys) būtų statistiškai prasmingos. Nes norima juk atvaizduoti gyventojų santykį su g y v e n a m u plotu. Taigi reikalinga, kad gyventojai būtų n u o l a t i n i a i to ploto gyventojai, o ne atsitiktinai kuria nors proga susirinkę. Be to, ir ploto duomenys turi būti statistiškai prasmingi, jei norima atskirų valstybių arba rajonų šios rūšies duomenis lyginti arba vienoj statistinėj lentelėj atvaizduoti. Jei Lietuvoj 1936 m. pradžiai Statistikos Metraštyje parodyta, kad 1 qkm tenka 45,6 gyventojai, o yra žinoma, kad, pav., Švedijoj ar kuriam kitam krašte 1 qkm tenka tik 9 gyventojai, tai dar nereiškia, kad Lietuva penkis kartus tirščiau apgyventa. Ir Lietuvoj 45,6 žmonės 1 qkm galima gauti tik išskyrus didesniųjų ežerų plotus. O juos priskyrus, gyventojų tirštumas išeitų mažesnis. Švedijoj gi tiek maža žmonių išeina 1 qkm tik dėl to, kad vandenų ir kitų negyvenamų plotų ten yra labai daug.

Ta aplinkybė, kad lyginiai skaičiai vaizduoja d v i e j ų statistinių eilučių arba s k i r t i n g ų vienos statistikos dalinių (pažymių) santykį, įgalina statistinius lyginius dvejopu būdu apibrėžti. Pav., užuot rodyti, kiek gyventojų tenka 1 qkm, galima parodyti, kaip didelis žemės plotas tenka 1 gyventojui. Parinkimas vieno ar kito metodo statistikos lyginiams pareina iš dalies nuo įsigalėjusių lyginimo papročių, iš dalies nuo pasistatytojo tikslo. Pav., įsigalėjusia statistikoj praktika gyventojų santykis su žemės plotu, paprastai, rodomas lyginiais skaičiais, kurie gaunami gyventojų bendrąjį skaičių padalinus ket-

virtainių kilometrų (qkm) skaičiumi $\frac{\text{gyventojai}}{\text{plotas}}$. Tačiau specialiniams uždaviniams aiškinti, k. a. sodžiaus gyventojų būklei nušviesti, dažnai būna tiekiami ir lyginiai skaičiai, kurie gaunami žemės ūkyj naudojamos žemės plotus (hektarus) padalinus sodžiaus gyventojų skaičiumi. Šiuo būdu gautieji statistiniai lyginiai parodo, kiek ha tenka vienam gyventojui. Taip pat pasiėmus lyginiais skaičiais vaizduoti švietimo būklę bet kuriame krašte, galima rodyti: a) kiek qkm tenka vienai mokyklai, ir b) kiek mokyklų išeina 1 qkm. Tiek vienos, tiek antros rūšies lyginiai skaičiai šiuo atveju bus vienodai prasmingi. Tačiau, jei pasiimtume gyventojų vaisingumą ar mirtingumą vaizduoti ne mirimus ar gimimus lygindami su gyventojų skaičiumi, o gyventojų skaičių su mirimais arba gimimais, tat pavaizdumas jau nukentėtų: viena, dėl to, kad statistikoj jau iš seno mirtingumo ir vaisingumo statistiniai lyginiai patiekiami santykyje su 1000 gyventojų (rašoma ir sakoma, kad 1000 gyventojų tenka *x* gimimų ir *y* mirimų), o, antra, statistinio vaizdavimo objektu šiuo atveju yra mirtingumo arba vaisingumo reiškiniai (mirimai arba gimimai), o ne gyventojai.

Lyginiai santykio skaičiai visais atvejais pareina betarpiškai nuo bazinių skaičių ypatybės. Jei baze pasirinktieji statistikos skaičiai yra netikri, arba nėra logiškai susieti su lyginamais statistiniais duomenimis, tai ir statistiniai santykio lyginiai savo vertės nustoja. Mirtingumą vaizduojantieji lyginiai skaičiai tik tiek yra tikri ir patikimi mirtingumo rodikliai, kiek yra tikri ir patikimi gyventojų skaičiaus duomenys atskiriems metams ar atskiriems rajonams. Pav., jei tikėti oficialinės statistikos šaltiniams, 1936 mt. bendrieji gimimų koeficientai buvo didesni miestuose, negu sodžiuose (24,5 prieš 24,1‰), o Kauno ir Klaipėdos miestuose 1.000 gyventojų turėjo tekti net 27,4 ir 28,7 gimimai prieš 24,5 ir 19,0 gimimų tų pačių pavadinimų apskrityse. Tačiau šitie duomenys prieštarauja ne tik visame pasaulyje pastebėtam reiškiniui (gimimų koeficientai yra mažesni miestuose), bet ir tų pačių miestų gimimų koeficientams iš to laikotarpio, kai gyventojų statistikos skaičiai patikimesni buvo. Pav., gyventojus surašant (1923 mt.) Kauno mieste 1.000 gyventojų teko tik 19 gimimų, sekančiais 1924 mt. 21 ir t. t., kai tais pačiais metais visoj Lietuvoj 1.000 gyv. teko

28 ir 29 gimimai. Taip pat ir vienoms jungtuvėms 1923/25 mt. Kaune vidutiniškai teko dar 2,5 gimimų, kai visoj Lietuvoj vienoms jungtuvėms 1923 mt. teko dar 4,0 ir 1924 mt. 3,6 gim. Tuo tarpu pastaraisiais 1935 ir 1936 mt., kai visoj Lietuvoj vienoms jungtuvėms vidutiniškai išeina 3,2 gimimai ir 1.000 gyv. apie 24 gim., Kauno miestui, kuriam oficialinė statistika gimimų priskiria apie 28 kiekvienam 1.000 gyventojų (vadinasi, žymiai daugiau), vienoms jungtuvėms tenka tik 1,6—1,7 gimimai. Čia prieštaravimas tiek aiškus, kad tolimesni aiškinimai ir nereikalingi. O kadangi gimimų ir vedybų registracija ir statistika pastaraisiais metais tvarkoma taip pat, kaip ir prieš 12—14 metų, tai tas prieštaravimas tegali pareiti tik nuo to, kad Kaune dabartiniu metu gyvena daug daugiau žmonių, negu oficialinė statistika rodo. Galima net apytikriai apie tą skaičių spėti. Nes jeigu pirmaisiais po surašymo metais parodytieji gimimų koeficientai (1.000 gyv.) buvo išskaičiuoti teisingai ir jie sudarė tik apie 20‰, ir jeigu per visą šį laiką Kauno mieste tie koeficientai būtų išsilaikę visai nesumažėję, tai Kaune turėtų būti ne 105 ar 107 tūkst. gyventojų, o apie 145 tūkst. (jeigu 1936 m. Kauno mieste 2 893 gimimai sudarė 1.000 gyventojų 20 gim., tai $\frac{2893 . 1000}{20}$ = 144650 gyventojų). Jeigu Kauno mieste per tą laiką bendrieji gimimų koeficientai sumažėjo tiek pat, kiek ir visoj Lietuvoj vidutiniškai, tai dabartiniu metu jie galėtų siekti ne daugiau kaip 18‰; bet tada Kauno miestui tektų apie 160 tūkst. gyventojų. Jei gi Kauno miesto gimimų koeficiento sumažėjimą matuoti 1 jungtuvėms tenkančių vedybų sumažėjimu, kuris viršija 30%, tai bendrieji koeficientai tesudarytų ne daugiau kaip 17‰; ir tada Kauno mieste turėtų būti apie 170 tūkst. gyventojų.

Šie pavyzdžiai rodo, kad statistinio santykio lyginiai gali turėti ir tokių ypatybių, kurios skaičiais vaizduojamą reiškinį iškreipia, ar net falsifikuoja. Klasišką tokios falsifikacijos pavyzdį patiekia prof. Wagemann savo „Narrenspiegel der Statistik" 151 psl., kuriam skaičiai paimti iš Westergaard'o „Mortalität und Morbilität" tyrinėjimų. Būtent, prof. Westergaard tyrinėjimais, Švedijoj 1881—90 laikotarpyje 1.000 gyventojų žemiau apibrėžtoms amžiaus grupėms teko mirimų:

22—26	amžiaus	grupei	6,13
27—31	„	„	5,89
32—36	„	„	5,88

Šie skaičiai, rodos, aiškiausiai įrodo (nes paimta net 10 metų vidurkiai, taigi galimos atsitiktinio pobūdžio paklaidos eliminuotos), kad pradedant nuo *22* metų ligi *36* metų amžiaus mirtingumas mažėja. Bet štai tie patys duomenys suskirstyti šeiminės padėties pažymiais:

Iš 1000 gyventojų pažymėtose amžiaus grupėse pasimirė:

amžiaus grupės	nevedę	vedę
22—26 m.	6,70	3,80
27—31 „	7,80	4,19
32—36 „	8,63	4,86

Čia patiektieji tie patys duomenys rodo visai priešingą vaizdą: tiek nevedusių, tiek vedusių senesnėse amžiaus grupėse mirtingumas didėja. Bet kaip tada suderinti šiuos duomenis su anksčiau patiektaisiais? Mįslė išsisprendžia visai paprastai: nevedusiųjų amžiaus grupėse mirtingumas beveik dvigubai didesnis kaip vedusiųjų; o kadangi jaunesniojo amžiaus grupėse nevedusiųjų proporcija žymiai didesnė, tai ji ir nulemia bendrųjų mirtingumo koeficientų dydį jaunesniojo amžiaus grupėse. Vėlesniojo amžiaus grupėse, nevedusiųjų proporcijai žymiai sumažėjus, sumažėja ir jų įtaka bendriesiems mirtingumo koeficientams. Savaime gi atskirose amžiaus fazėse, suskirsčius juos smulkesnių dydžių skaitomais periodais, mirtingumo mažėjimas yra galimas ir sutinkamas reiškinys (pav., bendrasis vaikų mirtingumas mažėja ligi 10—12 metų). Čia taip pat gali būti neliestas klausimas — dėlko *22—36* mt. amžiaus grupėse vedusiųjų ir nevedusiųjų mirtingumas taip pat žymiai skiriasi.

Aplamai imant, visų pirma tenka skirti: a) bendruosius arba brutto santykio lyginius nuo lyginių skaičių, kuriuos galima pavadinti — b) grynaisiais arba netto santykio statistiniais lyginiais. Be to, gali būti dar lyginiai santykio skaičiai ir kuriai nors mažiau ar daugiau griežtai apibrėžtai ma-

sės ar masinio reiškinio grupei. Šios rūšies lyginius santykio skaičius pavadinsime c) specifiniais statistinio santykio lyginiais.

Patogumo ir kaikuriais kitais sumetimais šios rūšies santykio skaičius sutarkime vadinti tiesiog statistiniais lyginiais (viensk. lyginys). Tuo būdu gausime: a) bendruosius (brutto), b) grynuosius (netto) ir c) specifinius statistikos lyginius.

Visi šios rūšies skaičiai gali būti pareiškiami trupmenomis, procentais, promiliais ar kuriais nors kitais skaičiais. Statistikos šaltiniuose, taip pat rašomojoj ir gyvojoj kalboj, kai kada šios rūšies skaičiai vadinami koeficientais, rodikliais, rodikliniais skaičiais ir kt.

Lyginių kategorijos priklauso ir vadinamieji indeksai. Bet kadangi indeksai ypač judėjimo statistikoj ypatingai svarbų vaidmenį vaidina ir jų sudarymo metodų yra labai daug, tai indeksų sudarymo ir jų pritaikymo problemas nagrinėsime vėliau visai atskirai.

a) Bendrieji (brutto) statistinio santykio lyginiai.

Paprastai statistikos šaltiniuose bet kurie reiškiniai yra vaizduojami bendraisiais (brutto) lyginiais. Pav., natūrinio gyventojų judėjimo (vedybų, gimimų, mirimų) lyginiai patiekiami promiliais (1000 gyv.), t. y. bendrieji vedybų, gimimų, mirimų ir natūrinio prieauglio skaičiai dalomi bendruoju gyventojų skaičiumi, o dalmuo dauginamas tūkstančiu. Jei bendrąjį gyventojų skaičių sutarsime žymėti raide N, o bendruosius gimimų, mirimų ir vedybų skaičius raidėmis G, M ir V, tai atitinkami lyginiai (L) promiliais bus gaunami pagal šias formulas:

$$\text{gimimams } L_g = \frac{G.1000}{N}$$

$$\text{mirimams } L_m = \frac{M.1000}{N}$$

$$\text{vedyboms } L_v = \frac{V.1000}{N}$$

$$\text{Natūriniam prieaug. } L_p = \frac{(G-M).1000}{N}$$

Statistikos šaltiniuose ir šiaip raštuose šios rūšies lyginiai dažniausiai vadinami natūrinio gyventojų judėjimo koeficientais. Panašiai gaunami ir nusikaltimų arba nusikaltėlių lyginiai kriminalinėj statistikoj. Tik baziniai duomenys dažniausiai čia prilyginami 10 000 ar net 100 000 gyv., o ne 100 ar 1000.

Aplamai, bet kurio reiškinio R, lyginant jį su bet kurios statistinės masės vienetų bendruoju skaičiumi N, bendrasis (brutto) lyginys L_r gaunamas pagal formulą:

$$L_r = \frac{R}{N} \qquad \text{(IV)}.$$

Jei sutariama statistinį lyginį žymėti procentiniame santykyj, promiliais ar kt., tai IV formula atitinkamai koriguojama, pav.:

$$L_r{}^{0}/_{0} = \frac{R.\,100}{N} \qquad \text{(IVa)}.$$

$$L_r{}^{0}/_{00} = \frac{R.\,1000}{N} \qquad \text{(IVb)}.$$

ir t. t.

Iš pirmo pažvelgimo gali atrodyti, kad šiuo būdu gautieji lyginiai visais atvejais yra statistiškai prasmingi bei pakankami bet kurios rūšies reiškiniams atvaizduoti bei lyginti: nes atsižvelgiama juk ne tik tų reiškinių (gimimų, mirimų ar kt.) absoliutinio skaičiaus, kurie gali pareiti ir nuo stebėtojo laiko skirtumų, bet ir nuo gyventojų bendrosios masės, kuri šios rūšies reiškinių absoliutinius skaičius daugiausia veikia ir nulemia. Tačiau jau paviršutinis samprotavimas parodys, kad šios rūšies statistiniai lyginiai ne visais atvejais gali parodyti charakteringas ir tikrenybę atitinkančias reiškinio ypatybes. Pav., imkime mirimus. Yra jau žinoma, kad pirmamečių kūdikių mirtingumas siekia ligi 15 ar net 20%. Pradedant antraisiais metais mirtingumas mažėja ligi maždaug 9—12 metų, paskui būna, palyginti, pastovus, arba pastebimas tik nedidelis mirtingumo didėjimas, ligi maždaug 40 metų. Po to mirtingumas ima didėti smarkiau; senių gi nuo 80 m. mirtingumas vėl pasiekia, ar net pralenkia, kūdikių mirtingumą. Tuo būdu išeina, kad juo kurioj šalyj ar tautoj yra daugiau vaikų ir senių, tuo

didesnis turi būti toj šalyj bendrasis mirusiųjų skaičius, o tuo pačiu didesni ir bendrieji mirimų lyginiai (bendrieji mirtingumo koeficientai). Ir atvirkščiai: bendrojoj gyventojų masėj, palyginti, nedidelės vaikų ir senių amžiaus grupės paveiks bendruosius mirimų skaičius, ir tuo pačiu bendrieji mirtingumo koeficientai išeis mažesni, — nežiūrint į tai, kad mirtingumas atskirai paimtose amžiaus grupėse gali būti didesnis. Taip pat bendrieji vedybų lyginiai (vedybų koeficientai) bus didesni ten, kur vadinamojo vedamojo amžiaus grupės yra gausingesnės, o senių ir vaikų amžiaus grupės — nedidelės. Net padaugėjus vedyboms nevedusiųjų amžiaus grupėse, bendrieji vedybų koeficientai gali išeiti mažesni, jei tuo pačiu laiku savo lyginamuoju svoriu būtų padidėjusios vaikų, senių ar apsivedusiųjų amžiaus grupės. Tos pačios arba panašios aplinkybės gali nulemti ir kitus bet kurios rūšies reiškinio statistinius santykio lyginius. Iš čia seka reikalas, norint kurį nors reiškinį nuodugniau ištirti ir tikrenybę atitinkančiais statistiniais santykio lyginiais atvaizduoti, — patiekti ir grynuosius st. santykio lyginius.

b) Grynieji (netto) statistinio santykio lyginiai.

Grynieji bet kurios masės ar masinio reiškinio lyginiai privalo atvaizduoti duotojo reiškinio charakteringas ypatybes arba kitimus santykyje tik su tais lyginiui imamos masės daliniais, kurie to reiškinio ypatybes arba kitimus nulemti gali. Vadinasi, tenka išskirti iš jos tuos dalinius arba grupes, kurios reiškinio susidaryme nedalyvauja. Pav., norint gauti vedybų grynuosius santykio lyginius, tenka išskirti nevedamojo amžiaus vaikus ir visus vedusius. Gimimų grynieji santykio lyginiai gali būti skaičiuojami dvejopu būdu: a) išskiriant vaikų amžiaus grupes ir b) gimimų skaičių imant santykyj su vad. gimdomojo amžiaus moterų skaičiumi (paprastai nuo 15 ligi 50 m.). Pastaruoju būdu išskaičiuotas lyginys vadinamas dar v a i s i n g u m o koeficientu.

Simboliais bendruosius ir grynuosius lyginius galima šitaip išreikšti.

Jei kuris nors reiškinys R masės vienetų bendrame skaičiuje N yra pasiskaidęs nevienodai, arba kai kuriose grupėse

n visai neatstovaujamas (pav., vedybų skaičius visai nepareina nuo vaikų skaičiaus), tai bendrasis lyginys $L_r = \frac{R}{N}$ palieka kaip buvęs. Tačiau grynasis st. lyginys turės būti koriguotas pagal formulą:

$$L'_r = \frac{R}{N-n} \qquad (V).$$

Šioj formuloj n turi atstovauti visus tuos bendrosios N masės dalių (grupių) vienetus, kurių R reiškinys nepriklauso. Kitaip tariant, ši formula rodo, kad sudarant grynuosius statistinius santykio lyginius, visuomet reikia išskirti vad. nedalyvaujančias masės dalis.

Grynųjų santykio lyginių skaičiavimo pavyzdžiams galima imti gimimų duomenis. Iš gyventojų surašymo duomenų statistikos šaltiniuose mes žinome, kad 1924 m. Lietuvoje buvo 2 170 616 gyv. tame skaičiuje 30,65% sudarė gyventojų amžiaus grupė 15 m. ir 69,35% nuo 15 m. ir senesni. Absoliutiniais skaičiais vaikai ir mažamečiai ligi 15 m. sudarė 665 392. Tai ir bus vad. nedalyvaujančioji masė, kurią grynajam gimimų lyginiui tenka išskirti. Jei mes bendrąjį gyventojų skaičių sutarsime žymėti N, vad. nedalyvaujančią masės dalį — n ir bendrąjį gimimų skaičių — G, tai bendrasis gimimų lyginys $L_g = \frac{G}{N}$, o grynasis $L'_g = \frac{G}{N-n}$. Norint lyginius skaičius gauti promiliais (1000 gyventojų), tenka gimimų skaičių — G padauginti 1000, ir tada atitinkamos formulos gausis:

$$L_{g\,0/00} = \frac{G.\,1000}{N} \quad \text{ir} \quad L'_{g\,0/00} = \frac{G.\,1000}{N-n}.$$

Iš tų pačių statistikos šaltinių žinodami, kad 1924 m. gimė 63 864, galime tuojau išskaičiuoti ir atitinkamus grynuosius gimimų statistinio santykio lyginius arba g r y n u o s i u s gimimų k o e f i c i e n t u s. Kai

$$G = 63\,864,\ N = 2\,170\,616,\ n = 665\,392,$$

tai

$$L_g = \frac{63\,864.1000}{2\,170\,616} = 29{,}34,$$

$$L'_g = \frac{63\,864.1000}{2\,170\,616 - 665\,392} = \frac{63\,864\,000}{1\,505\,224} = 42{,}43^0/_{00}.$$

Panašiu būdu gausime ir vadinamuosius vaisingumo koeficientus. Paėmę bendrąjį gimimų skaičių dalome vad. gimdomojo amžiaus moterų skaičiumi (atmetus jauniausiąją moterų amžiaus grupę ligi 15 m. ir seniausiąją nuo 50 m. ir daugiau).

Iš tų pačių statistikos šaltinių mes galime nustatyti, kad 1924 m. iš bendrojo moterų skaičiaus (1 138 884) moterų grupė nuo 15 ligi 50 m. sudaro 52,43% arba 597 117, o nedalyvaujančioji masė (ligi 15 m. ir nuo 50 m.) 47,57% arba 541 767. Tuo būdu vasingumo koeficientas gausis 106,95 promiliai, nes:

$$\frac{63\,864.1000}{1\,138\,884 - 541\,767} = \frac{63\,864\,000}{597\,117} = 106{,}95^0/_{00}.$$

Kad atskirtume vaisingumo koeficientą nuo grynojo gimimų lyginio, galima sutarti vaisingumo lyginį žymėti L''_v, bendrąjį moterų skaičių N' ir gimimuose „nedalyvaujančių" (ligi 15 m. ir nuo 50 m.) moterų skaičių n'. Tuo būdu formula V vaisingumo koeficientu promiliais pareikšta gaus variantą:

$$L''_{v}{}^0/_{00} = \frac{G.1000}{N' - n'} \qquad \text{(Va)}.$$

Panašiu būdu skaičiuojami grynieji vedybų, prasikalstamumo ir kt. statistinio santykio grynieji lyginiai arba koeficientai. Visais tokiais atvejais iš bendrosios masės išskiriami vad. nedalyvaujantieji, arba labai mažai tedalyvaujantieji masės daliniai (pav., vedybų arba prasikalstamumo grynieji koeficientai skaičiuojami išskyrus mažamečius ir p.). Svarbiausias uždavinys — tai kiek galint tiksliau tuos nedalyvaujančius ir „mažai" tedalyvaujančius masės dalinius išskirti.

Specialinę problemą sudaro grynųjų mirtingumo ir natūrinio prieauglio koeficientų skaičiavimas. Dalykas toks, kad atskirai paimtose amžiaus grupėse mirtingumas yra labai nevienodas. Pav., kai pirmamečių kūdikių mirtingumas siekia ligi 15% ir net geriausiu atveju tik apie 5%, tai antramečių kūdikių mirtingumas visur yra keleriopai mažesnis, vėliau nukrinta ligi keletos promilių ir tik senesnėse amžiaus grupėse

mirtingumas vėl ima didėti ir, pagaliau, maždaug nuo 75 m., prisiartina pirmamečių kūdikių mirtingumui.

Tuo būdu atskirų amžiaus grupių mirtingumo skaičių eilutė nesudaro statistinio vieneto. Dėlto ir bendrieji mirtingumo koeficientai, kai jie veikiami keletos statistinių dalinių nevienodai, savo griežtumo nustoja — nežiūrint į tai, kad šiuo atveju nedalyvaujančių statistinės masės dalinių visai nėra.

Statistiškai metodologiniu atžvilgiu tektų tokią statistinę masę padalinti į tiek dalinių, kiek yra statistinės dispersijos atžvilgiu skirtingų mirtingumo grupių, ir kiekvienai tokiai grupei išskaičiuoti specifinius, tik ją atitinkančius, mirtingumo lyginius. Tačiau iš mirtingumo statistikos reikalaujama, kad ji patiektų statistiškai prasmingus mirtingumo koeficientus visai gyventojų masei tokio pat patikimumo, kaip grynieji gimimų bei vaisingumo, vedybų ar prasikalstamumo lyginiai.

Deja, čia jau atsiranda sunkiai nugalimų kliūčių. Visų pirma, bet kurių metų bendrąjį mirusiųjų skaičių nulemia gyvųjų amžiaus grupės iš labai ilgo, ligi 100 metų siekiančio, periodo, ir tik mirusiųjų mažumą nulemia pirmamečių ir mažamečių masės. Norint gauti statistiškai prasmingus grynuosius, t. y. savo rūšies svertinius mirtingumo koeficientus, tektų kiekvienų metų mirusiųjų amžiaus grupes „atsverti" tų metų gyvųjų amžiaus grupių svorio koeficientais. Tačiau tokius svorio koeficientus tegalima gauti tik tiesioginio surašymo keliu. Gyventojų gi surašinėjimai niekur kas metai nedaromi, o kai kada praeina keliolika ir net keliasdešimt metų tarp dviejų tokių surašymų. Tuo būdu ir tikrosios amžiaus grupių sudėties nė juos atitinkančių svorio koeficientų visiems atskirai paimtiems metams gauti negalima. Be to, gali atsitikti, kad kurių nors nenormalinių sąlygų įtakoj susidariusi anormalinė amžiaus grupė, jei tos grupės specifinis mirtingumo koeficientas griežtai skiriasi nuo normalinio mirtingumo koeficiento, gali žymiai paveikti, t. y. iškreipti ir bendrąjį mirtingumo koeficientą (pav., jei būtų nepaprastai didelės arba nepaprastai mažos vaikų arba senių amžiaus grupės, arba nepaprastai didelė arba nepaprastai maža amžiaus grupė nuo 6 ligi 15 metų, kur mirtingumas aplamai yra mažiausias).

Šios rūšies sunkumams nugalėti ir net kiek patikimesnius mirtingumo koeficientus gauti mėginama bendruosius mirtin-

gumo lyginius koriguoti pasirėmus vadinama „normaline“ gyventojų amžiaus grupių sudėtimi. Šitokia normalinė amžiaus grupių sudėtis, manoma, geriausiai yra atstovaujama amžiaus dalinių, sudarytų pasirėmus naujausiojo gyventojų surašymo duomenimis ir korigavus arba „išlyginus“ tuos duomenis senesniųjų surašymų duomenimis tiek, kiek senesnėsės amžiaus grupės gali dar veikti bendruosius mirimų duomenis.

Pavyzdžiui galima paimti mirusiųjų duomenis 1936 m. Tais metais Lietuvoj pasimirė 33 440 asmenys. Kadangi metų pradžioje pagal CSB duomenis būta beveik 2,5 mil. gyventojų, tai bendrasis mirimų lyginys pagal mums pažįstamą formulą $L_m = \frac{M. 1000}{N} = \frac{33\,440}{2\,500\,000} = 13{,}4^0/_{00}$. Tačiau žinome, kad gyventojų bendroji masė mirtingumo atžvilgiu taip pat sudaro nevienareikšmius dalinius. Skaičiavimui palengvinti sutarkime tik tris tokius dalinius imti: 1) iš pirmamečių kūdikių, 2) iš gyventojų nuo 1 ligi 59 metų amžiaus ir 3) iš likusiųjų gyventojų nuo 60 m. amžiaus. Jei šitų trijų dalinių lyginamąjį svorį paimsime iš 1923 m. visuotiniojo surašymo duomenų, tai gausime, apvaliais skaičiais, šitokią lentelę:

1.	Amžiaus grupė	nuo	0— 1	sudarė	2,4%
2.	„	„	1—60	„	87,1%
3.	„	„	60 ir daugiau		10,5%

Bet tenka dar atsižvelgti, kad 1936 m. mirusiųjų skaičių nulėmė ne tik amžiaus grupių sudėtis, konstatuota per 1923 m. surašymą, bet ir amžiaus grupės iš senesniųjų metų. O, pav., 1897 m. surašymo duomenimis tų pačių amžiaus grupių lyginamasis svoris pasireiškė šitaip (apytikriai):

amžiaus grupės	% %
0— 1	3,0
1—60	89,2
60 ir daugiau	7,8

Tačiau ir 1897 m. amžiaus grupių svorio koeficientai negali būti laikomi normaliais — jau vien dėl to, kad prieš 40 metų gimusiųjų amžiaus grupės karo metu buvo labai sužalotos. Turint galvoje, kad kiekvienoj visuomenėj yra bent treje-

tas gentkarčių, su atitinkama jauniausios kartos persvara, tikslingiausia būtų ir „normalinių“ amžiaus grupių koeficientus sudaryti pasirėmus kiekvienos kartos amžiaus grupių sudėtimi maždaug tokiais atstumais:

a) einamųjų metų (pagal naujausį surašymą) amžiaus grupių sudėtis;

b) prieš 20—30 metų buvusios amžiaus grupės;

c) prieš 60 metų buvusios amžiaus grupės.

Visų trijų (a+b+c) variantų vidurkis duotų „normalinę“ amžiaus grupių sudėtį. Ši „normalinė“ amžiaus grupių sudėtis galėtų būti pagrindu svertiniam mirtingumo koeficientui skaičiuoti.

Dar vienas būdas „pataisytiems“ arba svertiniams mirtingumo koeficientams gauti — tai svorio koeficientus paimti iš vad. mirtingumo lentelės, kuri paprastai sudaroma pasikliaujant naujausiojo gyventojų surašymo duomenims amžiaus grupėms, ir juos panaudoti mirtingumo koeficientams sudaryti.

Aplamai, jei mes sutarsime bendrąjį bet kurių metų mirusiųjų skaičių žymėti raide M, atskirų amžiaus grupių mirusius $m_1, m_2, \ldots . m_n$ ir bendrojo gyventojų skaičiaus N atitinkančias amžiaus grupes $n_1, n_2, \ldots, n_n$, ir jų „normalinės sudėties“ statistinius dalinius raidėmis $d_1, d_2 \ldots ., d_n$, tai koriguotas arba „grynasis“ mirtingumo lyginys bus gaunamas pagal formulą:

$$L'_m = \frac{\frac{m_1 . d_1}{n_1} + \frac{m_2 . d_2}{n_2} + \ldots\ldots + \frac{m_n . d_n}{n_n}}{d_1 + d_2 + \ldots\ldots + d_n} = \frac{\Sigma \frac{m_i . d_i}{n_i}}{\Sigma d_i} \qquad \text{(VI.}$$

Jei sutarsime atskirų amžiaus grupių mirtingumo koeficientus $\frac{m_i}{n_i} = q_i$ žymėti raidėmis $q_1, q_2, \ldots, q_n$ aplamai q_i tai VI formula gaus variantą:

$$L'_m = \frac{q_1 d_1 + q_2 d_2 + \ldots . + q_n d_n}{d_1 + d_2 + \ldots . + d_n} = \frac{\Sigma q_i d_i}{\Sigma d_i} \qquad \text{(VIa).}$$

Formulos rodo, kad faktinieji arba specifiniai kiekvienos, skirtingą mirtingumą turinčios, amžiaus grupės mirimų koeficientai dauginami tariamai „normalinės“ amžiaus grupių sudėties daliniais, kurie šiuo atveju atstoja svorio koeficientus. Tuo būdu, gautąją sumą padalinus dalinių suma, gaunamas grynasis mirtingumo lyginys.

Kad paaiškėtų šitų vad. „grynųjų“ arba koriguotųjų mirtingumo koeficientų statistinė prasmė, pavyzdžiui paimkime tuos pačius 1936 m. mirimų duomenis. Jei sutarsime 1923 m. surašymo amžiaus grupių dalinius laikyti atitinkančiais ir 1936 m. dalinius, o 1897 m. surašymo konstatuotus dalinius — „normaliniais“, tai gausime šitokį vaizdą:

$M=33\,440$; $m_1=7\,755$, $m_2=11\,278$ ir $m_3=14\,407$.
$N=2\,500\,000$; $n_1=60\,000$, $n_2=2\,177\,500$ ir $n_3=252\,500$.

Tuo būdu gauname q_i:

$q_1=7\,755:60\,000(=)0{,}129$,
$q_2=11\,278:2\,177\,500(=)0{,}005$,
$q_3=14\,407:252\,500(=)0{,}074$.

Kai sutarėme „normaliniais“ amžiaus grupių daliniais d_i laikyti 1897 m. gyventojų surašymo dalinius

$$d_1=3{,}0,\quad d_2=89{,}2,\quad d_3=7{,}8 \text{ ir } \Sigma d_i=100,$$

tai formula:

$$L'_m=\frac{\Sigma q_i d_i}{\Sigma d_i},$$

simbolius pakeitus skaičiais, duos mirtingumo lyginį

$$L'_m=\frac{0{,}129.3{,}0+0{,}005.89{,}2+0{,}074.7{,}8}{3{,}0+89{,}2+7{,}8}=$$

$$=\frac{0{,}388+0{,}462+0{,}428}{100}=\frac{1{,}278}{100}=$$

$$=0{,}01278,$$ t. y. 12,8 promiliai (apytikriai).

Šitaip išskaičiuotas mirtingumo lyginys išeina žymiai mažesnis už bendrąjį Lm, kuris siekia 13,4%. Tačiau jau bus kitas vaizdas, jei normaliniais amžiaus grupių daliniais d_i laikysime ne 1923 m. surašymo dalinius, o atitinkamų amžiaus grupių mirtingumo koeficientus skaičiuosime pagal 1897 m. duomenis. Šiuo atveju turėsime:

$$n_1=75\,000\left(\frac{2\,500\,000.3}{100}=75\,000\right),$$

$n_2=2\,230\,000$,
$n_3=195\,000$.

Tuo būdu:

$q_1 = 7\,755 : 75\,000 = 0{,}103$
$q_2 = 11\,278 : 2\,230\,000 = 0{,}005$ ir
$q_3 = 14\,407 : 195\,000 = 0{,}074.$

O kadangi tariama, kad 1936 m. amžiaus grupių sudėtis atitiko 1923 m. surašymo konstatuotą sudėtį, taigi $d_1 = 2{,}4$, $d_2 = 87{,}1$ ir $d_3 = 10{,}5$, tai ta pati VI-ta formula duos:

$$L''_m = \frac{0{,}103.2{,}4 + 0{,}005.87{,}1 + 0{,}074.10{,}5}{2{,}4 + 87{,}1 + 10{,}5} =$$

$$= \frac{0{,}248 + 0{,}440 + 0{,}776}{100} = \frac{1{,}464}{100} =$$

$$= 0{,}01464 \text{ arba } 14{,}64\,^0/_{00}.$$

Šie pavyzdžiai rodo, kad vad. grynieji arba koriguotieji mirtingumo lyginiai gali būti didesni arba mažesni už bendruosius — pagal tai, kokią amžiaus grupių sudėtį laikysime normaline. Aplamai, tačiau, griežtai objektyviniais statistikos duomenimis pasirėmus nustatyti vad. „normalinius" amžiaus grupių dalinius nėra galima, ir mažesnis ar didesnis dirbtinumas čia neišvengiamas. Taip pat nelengva surasti ir grynuosius mirtingumo koeficientus. Pav., jei paanalizuoti kad ir 1936 m. mirusiųjų ir to meto visų gyventojų amžiaus grupių sudėtį, tai, visų pirma, nerasime pagrindo pirmamečių grupės skaičių vertinti 75 000: nes 1934 m. iš viso gimė 60 770 kūdikiai, 1935 m. — 57 970, o 1936 m. — 60 446. O kadangi 1934 m. pirmamečių kūdikių išmirė 10 059 ir 1935 m. 7 146, tai ir 1936 m. pradžioj pirmamečių negalėjo būti daugiau kaip 60 000, t. y. tik apie 2,4% iš bendrojo gyventojų skaičiaus, o ne 3,0%. Iš kitos pusės yra pagrindo teigti, kad antroji amžiaus grupė tarp 1—59 m., lyginant su 1923 m., bus kiek didesnė, nes šios grupės karo metų sužalojimų dalis papildyta iš 1922—1936 m. laikotarpio, t. y. 14 metų, gimusiais. Be to, karo sužalotųjų amžiaus grupių dalis (visi, kurie 1923 m. turėjo nemažiau kaip 46 m.) išmirė arba perėjo į paskutinę, 60 metų ir senesniųjų grupę. Ši grupė, apytikriais skaičiavimais, galėjo sudaryti apie 8,6%, t. y. 215 000. Bet jeigu $n_1 = 60\,000$, o $n_2 = 215\,000$, tai $n_2 = 2\,225\,000$ (nes kai $n_1 + n_2 + n_3 = 2\,500\,000$, tai $n_2 = 2\,500\,000 - (60.000 + 2\,150\,000) =$

$=2\,225\,000$. Taigi šiose aplinkybėse faktinieji mirtingumo koeficientai q_i jau šitaip atrodys:

$$q_1=7\,755:60\,000(=)0{,}129$$
$$q_2=11\,278:2\,225\,000(=)0{,}005$$
$$q_3=14\,407:215\,000(=)0{,}067$$

Dabar tik klausimas, kokius dydžius reikia laikyti „normaliniais" amžiaus grupių daliniais (d_i)?

Ieškant į šį klausimą atsakymo tenka pripažinti faktą, kad ir Lietuvoje gimimų koeficentai mažėja, o pastovesni yra absoliutiniai gimimų skaičiai, kurie jau pradeda nebepasiekti 60 000 ribos. Šitose aplinkybėse, kai bendrasis gyventojų skaičius vis dar apie 1% kas metai priauga, numatomoj ateityj pirmamečių grupė greičiau lygiuosis prie statistinio dalinio maždaug apie 2%. Iš kitos pusės pastebimas vidutinio gyventojų amžiaus ilgėjimas. Vadinasi, senių grupė nuo 60 metų, net ir veikiama karo meto sužalojimų, galės sudaryti maždaug 9%. Taigi 1—60 m. amžiaus grupę atstovaujantis dalinys sudarytų apie 89%. Bet jeigu normaliniais amžiaus grupių daliniais paimsime $d_1=2\%$, $d_2=89\%$, $d_3=9\%$, tai

$$L'''_m=\frac{0{,}129.2+0{,}005.89+0{,}067.9}{2+89+9}=$$
$$=\frac{0{,}258+0{,}445+0{,}603}{100}=\frac{1{,}306}{100}$$
$$=0{,}01306 \text{ arba } 13{,}06\%.$$

Kaip matome, vadinamų „grynųjų" mirtingumo lyginių ieškojimas yra ir sudėtingas, ir problematiškas dalykas. Kai kurių autorių (pav. Winkler'io) patarimas — grynąjį mirtingumo lyginį skaičiuoti iš mirtingumo lentelės vienaip ar kitaip sudarytais koeficientais, visų problemų taip pat neišsprendžia (pati lentelė taip pat tėra tik „apytikrė" ir neapsaugota nuo ateities kitimų).

Aplamai, vad. grynųjų statistikos lyginių skaičiavimas statistiškai prasmingas visais atvejais yra tik tuomet, kai kuri nors statistinės masės dalis analizuojamo ieškinio susidaryme nedalyvauja[1].

[1] Pačiu naujausiuoju laiku gyventojų natūrinio judėjimo statistikoj pradedami naudoti vad. reprodukcijos koeficientai, kurie taip

c) Specifiniai statistinio santykio lyginiai.

Kai grynieji santykio lyginiai turi atvaizduoti bet kurio reiškinio saitus su visa mase, išskyrus tos masės vad. nedalyvaujančias ar menkai tedalyvaujančias mases, tai specifiniai santykio lyginiai turi parodyti bet kurio reiškinio saitus su atskiromis masės dalimis.

Pavyzdžiui galime imti vėl tuos pačius gimimų ar kt. duomenis. Mes žinome ne tik tai, kad gimimų bendrieji skaičiai pareina nuo abiejų lyčių vaikų ir senesniojo amžiaus moterų skaičiaus, bet ir tai, kad vad. gimdomojo amžiaus atskirose grupėse gimimų gausumas yra nevienodas. Pav. 1935 mt. ir Lietuvoj 77% visų gimusiųjų tenka šeimoms, kur tėvas yra nesenesnis kaip 40 mt. ir net 93% gimimų tenka motinoms, nesenesnėms kaip 40 mt. Bet ir smulkesnėse tėvų amžiaus grupėse gimimai nevienodi. Vadinasi, galima ir pravartu tam tikrais atvejais žinoti ir specifinius gimimų lyginius atskiroms tėvų amžiaus grupėms, arba specifinius vaisingumo koeficientus moterų amžiaus grupėms. Panašiu būdu gaunami ir specifiniai mirtingumo lyginiai arba koeficientai vad. mirtingumo lentelėms. Taip pat galima bet kurio reiškinio specifinius lyginius gauti atskiroms tautybėms, tikyboms, verslui arba profesijoms ir t. t.

Jei kuris nors reiškinys R masėje N yra pasiskaidęs mažiau ar daugiau įvairiu gausumu, tai specifinių lyginių l_i bus tiek, kiek turėsime masės skirtingų grupių. Jei masės skirtingųjų grupių skaičius sutarsime žymėti a_1, n_2, n_3, ... n_n, o atitinkamus reiškinius r_1, r_2, r_3, ... r_n, tai specifiniai lyginiai l_i gaus šias formas:

---

pat dar skirstomi į bendruosius ir grynuosius reprodukcijos koeficientus. Šios rūšies st. lyginiai turi parodyti, ar duotojo meto gimimai bei natūrinis prieauglis yra pakankamas, kad gyventojų esamas skaičius ateityje nesumažėtų, ar nepakankamas. Metodai šiems gyv. reprodukcijos koeficientams gauti naudojami įvairūs. Tačiau visi sutaria, kad koeficientų skaičiavimo pagrindą sudaro gimdomojo amžiaus moterys (tarp 16—45 ar 15—50 mt.) ir metinis gimimų skaičius. Dėl kitų reprodukcijos koeficientų formulai reikalingų komponentų naudojimo nuomonės nesutaria (pav., imami dar natūrinio prieauglio koeficientai, gimdomojo amžiaus moterų specifiniai mirtingumo koeficientai, jų gyventino amžiaus tikimybiniai dydžiai ir t. t.).

Apie šio klausimo esmę ir metodus, taip pat literatūrą, suglaustoj formoj galima rasti prof. W. Winkler'io referate (Dr. *W. Winkler,* Einige alte und neue Masse des natürlichen Bevölkerungswachstums, Revue de l'Institut International de Statistique, 1938, I, pp. 25—49).

$$l_1 = \frac{r_1}{n_1},\ l_2 = \frac{r_2}{n_2},\ l_3 = \frac{r_3}{n_3}, \ldots,\ l_n = \frac{r_n}{n_n}.$$

Taigi, aplamai, bet kurio reiškinio r_i specifinis bet kurios grupės n_i lyginys l_i gaunamas pagal formulą:

$$l_i = \frac{r_i}{n_i} \qquad \text{(VII)}$$

Nesunku pastebėti, kad specifinių st. lyginių komponentai yra kartu ir bendrojo lyginio komponentai, nes:

$$R = r_1 + r_2 + \ldots + r_n = \Sigma r_i,\ o$$
$$N = n_1 + n_2 + \ldots + n_n = \Sigma n_i.$$

O bendrasis bet kurio reiškinio st. lyginys L_r pagal IV formulą:

$$L_2 = \frac{R}{N} = \frac{r_1 + r_2 + \ldots + r_n}{n_1 + n_2 + \ldots + n_n} = \frac{\Sigma r_i}{\Sigma n_i} \qquad \text{(IVc)}$$

Savaime, ir specifiniai st. santykio lyginiai l_i, lygiai kaip ir bendrieji, gali būti pareikšti procentais, promiliais ar kt. Tokiais atvejais ir VII formula atitinkamai pakeičiama, pav.:

$$l_{i\%} = \frac{r_i.100}{n_i}, \qquad \text{(VIIa)}$$

$$l_{i‰} = \frac{r_i.1000}{n_i},\ \text{ir t. t.} \qquad \text{(VIIb)}$$

Pavyzdžiui patiekiami du variantai specifinių mirtingumo koeficientų atskiroms amžiaus grupėms 1924 mt. Iš jų matyti, kad specifinių lyginių statistinis prasmingumas tik tuo atveju garantuojamas, jei juos sudarant atsižvelgiama charakteringųjų tam tikro reiškinio p a s i s k a i d y m o ypatybių. Pav., specifinis mirtingumo koeficientas 0—10 m. amžiaus grupei išeina 39,63‰. Bet iš lentelės 11 mes matome, kad pirmamečių mirtingumas 1924 m. siekė net 214‰, t. y. buvo didesnis kaip senių virš 80 mt. Tuo tarpu to paties pirmojo dešimtmečio 5—10 m. amžiaus grupėj mirtingumas tesudaro tik 3‰, t. y. nedaug kuo skiriasi nuo mirtingumo amžiaus grupėse tarp 15—30 metų.

Panašiu būdu gali būti randami gimimų, vedybų ir kt. specifiniai lyginiai, taip pat statistikos lyginiai atskirai miestams ir sodžiams, arba atskirai įvairioms tikyboms, tautybėms, verslams ir t. t.

Lentelė 11.

LIETUVOS GYVENTOJAI 1924 MT. IR MIRUSIEJI AMŽIAUS GRUPĖMIS.

Nr.	Amžius grupės (k_i)	Gyventojai (n_i)	Mirusieji (m_i)	Mirtingumas $\left(l_{o/oo} = \frac{m_i\,1000}{n_i}\right)$
1	0	54 095	11 556	213,62
2	1— 4	182 681	4 074	22,30
3	5— 9	182 016	968	5,32
4	10—14	261 163	661	2,53
5	15—19	270 474	874	3,23
6	20—24	237 219	978	4,12
7	25—29	186 228	911	4,89
8	30—39	241 652	1 486	6,15
9	40—49	193 766	1 616	8,34
10	50—59	176 030	2 516	14,29
11	60—69	141 888	4 101	28,90
12	70—79	62 519	3 973	63,55
13	80 ir d.	27 269	3 465	127,07
	Iš viso	2 217 000	37 179	

Lentelė 12.

LIETUVOS GYVENTOJAI 1924 MT. IR MIRUSIEJI AMŽIAUS GRUPĖMIS.

Nr.	Amžius grupės (k_i)	Gyventojai (n_i)	Mirusieji (m_i)	Mirtingumas $\left(l_{o/oo} = \frac{m_i\,1000}{n_i}\right)$
1	0— 9	418 792	16 598	39,63
2	10—19	531 637	1 535	2,89
3	20—29	423 447	1 889	4,46
4	30—39	241 652	1 486	6,15
5	40—49	193 766	1 616	8,34
6	50—59	176 030	2 516	14,29
7	60—69	141 888	4 101	28,90
8	70—79	62 519	3 973	63,55
9	80 ir d.	27 269	3 465	127,07
	Iš viso	2 217 000	37 179	

B. Vidurkiai ir st. eilučių dispersija.

5. Statistikos vidurkiai.

Gyvojoj žmonių kalboj daugelis gyvenimo reiškinių apibrėžiama tik apytikriai, t. y. sulig dažniau sutinkamais pavyzdžiais, tipais, normomis. Pav., jei kas nors, klausiamas, atsako, kad į darbo vietą ar tarnybą eina 15 min. ar pusę valandos, tai turi galvoje ne griežtai apibrėžtą laiką, per kurį jis kas dieną nueina, bet laiką, per kurį jis nueina dažniausia, t. y. nei perdaug skubindamas, nei stoviniaudamas. Ta pačia prasme kalbame apie vidutinį darbininką, vidutinį studentą, vidutinę kainą, vidutinį darbo laiką ir t. t.

Tiems patiems ar kitiems gyvenimo reiškiniams apibūdinti bei griežčiau apibrėžti statistika turi savuosius vardus bei metodus. Berods, dažniausias, arba žmonių kalbos vidutinis dydis yra pripažintas ir statistikoj, tik vadinamas jis čia dominantu arba (dažniau anglų) modusu. Tikrasis gi statistikos vidurkis yra vad. aritmetinis vidurkis, be to, medianas, t. y. pats vidurinis dydis.

Dalykas toks, kad statistikos vidurkis turi atstovauti visą statistinę eilutę, t. y. masės arba masinio reiškinio visumą. Bet patsai vienas aritmetinis vidurkis tam tikrais atvejais gali būti pačiai eilutei visai svetimas dydis. Pav., jei kur nors būtų 10 ūkių po 100 ha ir 90 ūkių po 10 ha, tai aritmetinis vidurkis išeitų 19 ha, nes:

$$\frac{10.100+90.10}{10+90}=\frac{1\,900}{100}=19.$$

Taigi gautasis vidurkis ne tik sau lygaus, bet nė panašaus arba artimo visoj ūkių masėj neturėtų. Tuo tarpu dominantas (dažniausias dydis) šitokioj eilutėj atstovautų 90% visos ūkių masės. Taip pat ir medianas, kuris yra pats vidurinis statistinės eilutės dydis, šiuo atveju atitiktų dominantą, vadinasi, taip pat atstovautų 90% visų ūkių.

Šie pavyzdžiai rodo, kad aritmetinis vidurkis ir medianas, arba medianas ir dominantas, tam tikrais atvejais gali būti reikalingi turėti šalimais vienas antro. Patsai vienas aritmetinis vidurkis, arba patsai vienas medianas ar dominantas dar

ne visa, kas reikalinga, apie statistinę eilutę pasako. Pavyzdžiui imkime skaičius 20, 25 ir 30, kurie gali pasikartoti nevienodai, sakysime: po 20 metų yra 10 studentų po 25—200 ir po 30 mt. tik 90 studentų. Jei norėsime sužinoti vidutinį studento amžių, tai paprastas ir svertinis aritmetiniai vidurkiai nesutaps, nes:

$$\frac{20+25+30}{3}=25, \quad \text{o}$$

$$\frac{20\,.\,10+25\,.\,200+30\,.\,90}{10+200+90}=$$

$$=\frac{200+5\,000+2\,700}{300}=\frac{7\,900}{300}=26{,}33\ldots.$$

Dar didesnis išeitų skirtumas, jei jauniausio amžiaus (po 20 mt.) studentų būtų daugiausia, t. y. 200, 25 mt. 90 ir 30 mt. tik 10. Tada svertinis vidurkis išeitų tik 21,83, nes:

$$\frac{20\,.\,200+25\,.\,90+30\,.\,10}{200+90+10}=$$

$$=\frac{4\,000+2\,250+300}{300}=$$

$$=\frac{6\,550}{300}=21{,}83\ldots$$

Tačiau jeigu kiekvienoj amžiaus grupėj studentų būtų po lygiai, tarsime po 100, tai ir svertinis jų amžiaus vidurkis būtų lygiai toks pat, kaip ir paprastas aritmetinis, nes:

$$\frac{20\,.\,100+25\,.\,100+30\,.\,100}{100+100+100}=$$

$$=\frac{7\,500}{300}=25, \text{ ir}$$

$$\frac{20+25+30}{3}=25.$$

Šie pavyzdžiai rodo, kad svertinis aritmetinis vidurkis nuo paprastojo vidurkio skiriasi tik tuo atveju, kai skirtingi dydžiai arba varijantai masėje atstovaujami nevienodai (vie-

nų yra daugiau, o kitų mažiau). O patsai aritmetinio vidurkio skirtumas išeis juo didesnis, kuo didesni patys vie etų (variantų) skirtumai ir kuo nevienodesnis jų pasiskaidymas.

a) Aritmetinių vidurkių apibendrinimai bei formulos.

Imkim bet kurią statistinę masę ar masinių reiškinių visumą. Jei sutarsime bet kurią statistinę masę žymėti X ir jos skaitomuosius vienetus žymėsime x_1, x_2, x_3, ... x_n, arba, aplamai, bet kurį skaitomąjį vienetą x_i, tai jų visumą, pareikštą raidę x, atstovaus skaičius n, kuris šiuo atveju atitiks paskutiniojo nario eilės numerį (x_n). Pav., jei masėj X tebūtų tik penketas vienetų x_i, tai šitie x_i dydžiai turėtų šiuos eilutės sekos numerius:

$$x_1, x_2, x_3, x_4, x_5.$$

Šiose aplinkybėse:

$$n = x_1 + x_2 + x_3 + x_4 + x_5 = \Sigma x_i,$$

arba

$$5x_i = X.$$

Patiektas pavyzdys rodo, kad jeigu tik skaitomieji x-inės ar y-inės eilutės nariai, arba varijantai (x_i arba y_i) yra s k a i t o m i e j i v i e n e t a i, tai X (arba Y) santykyje su n sudaro tiekybes (X = n), vadinasi, ir lygtyse bei formulose gali viena antrą pavaduoti. Tik tais atvejais, kai statistinės masės x_i vienetai eilutės sekoj pradedami žymėti 0 (x_0, x_1, x_2, ... x_i, ... x_n) paskutiniojo nario (x_n) eilinis skaičius (n) bus vienetu mažesnis už X, taigi, X = n + 1 = N.

Patį aritmetinį vidurkį, kuris gaunamas statistinės X eilutės x_i dydžių bendrąją sumą padalinus vienetų skaičiumi (n), sutarsime žymėti $\bar{x}$ (arba $\bar{y}$, $\bar{z}$ arba kt.).

Aplamai, x-inės eilutės aritmetiniam vidurkiui x skaičiuoti galioja ši formula:

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{1}{n} \Sigma x_i, \text{ t. y.} \qquad \text{(VIII).}$$

$$\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n}, \text{ arba}$$

$$\bar{X} = \frac{x_0 + x_1 + x_2 + \ldots + x_n}{n+1} = \frac{1}{N} \Sigma x_i .$$

Ši formula galioja visiems aritmetinio vidurkio atsitikimams. Tačiau, kai bet kuris x-inės eilutės variantas x_i visoje X masėje pasikartoja daugelį kartų (pav., mūsų anksčiau paimtame pavyzdyje iš šimto ūkių — 10 ūkių turėjo po 100 ha, o 90 — po 10 ha), nebūtų prasminga visos masės vienetus kiekvieną atskirai eilutės nr. žymėti ir iš eilės vis po vieną pridėti, kad visos X masės bendrąjį vienetų (x_i) skaičių (n) gauti.

Iš aritmetikos elementoriaus mes žinome, kad jeigu kuris nors x^i dydis pasikartoja, tai pakanka jį padauginti pasikartojimo skaičiumi, ir bus gauta bendroji visų dydžių suma. Jeigu x-nėj eilutėj vieni dydžiai pasikartoja dažniau, kiti — rečiau, tai vienodo dydžio vienetai arba variantai iš karto sudaro grupes x_i, prie kurių privalo būti pažyminys arba skaičius, rodąs šio x_i vienetų pasikartojimo skaičių arba gausumą.

Jei sutarsime x-inės eilutės bet kurio vieneto gausumą atstovaujančius variantus žymėti koordinatų sistemoj įprastu simboliu y, tai galėsime rašyti:

$$X = x_1 y_1 + x_2 y_2 + \ldots + x_i y_i + \ldots + x_n y_n = \Sigma x_i y_i .$$

Iš to seka, kad ir aritmetiniam vidurkiui formula VIII privalo naujo varianto:

$$\bar{X} = \frac{\Sigma x_i y_i}{\Sigma y_i} = \frac{1}{n} \Sigma x_i y_i , \text{ nes} \qquad \text{(VIIIa)}$$

$$\bar{X} = \frac{x_1 y_1 + x_2 y_2 + \ldots + x_i y_i + \ldots + x_n y_n}{y_1 + y_2 + \ldots + y_i + \ldots y_n} =$$

$$= \frac{\Sigma x_i y_i}{\Sigma y_i} = \frac{\Sigma x_i y_i}{n} .$$

Šiuo atveju, kaip matome, x-inės eilutės vienetų skaičių (n) atstovauja x_i variantų dažnumo (pasikartojimo) skaičių y_i bendroji suma Σy_i, taigi

$$n = \Sigma y_i .$$

Pagal VIIIa formulą išskaičiuotas vidurkis x vadinamas s v e r t i n i u aritmetiniu vidurkiu.

b) Vidurkiai ir jų suradimą palengvinantieji skaičiavimo metodai.

Aplamai, galima sakyti, kad nedidelės statistinės eilutės, arba negausingi statistinio stebėjimo objektai, vidurkiais vaizdavimo ir nereikalingi. Pav., jei mes turime tik 100 ūkių ir žinome, kad jų tarpe yra 90 ūkių po 10 ha ir 10 po 100 ha, tai daugiau mums nieko ir nereikia. Bet kita padėtis, kai ūkių ar kitų kurių statistikos skaitomųjų vienetų yra ne šimtai, o

šimtai tūkstančių, arba kai statistikos skaitomieji vienetai yra gauti reprezentaciniu metodu, vadinasi, iš keleto dešimčių ar šimtų turimųjų vienetų tenka spręsti apie tūkstančius. Šiuo atveju masę atstovaujančios statistinės eilutės analizis gali būti reikalingas ne tik visokiariopų vidurkių, bet ir kiek galint tikslesnių bei griežtesnių, o tuo pačiu laiku ir paprastesnių, jų suradimo bei skaičiavimo metodų.

Statistikos praktikoj aritmetinius vidurkius skaičiuojant visuomet svarbu kiek galint mažiau laiko sugaišti. Pasiekiama tai gali būti dviem būdais: a) turimus skaičius apvalinant ir b) parinkti pastovų dydį (konstantą) ir, vidurkį skaičiuojant, tik skirtumais operuoti.

Turimus skaičius apvalinant, gaunami apytikriai x_i skaičiai ir, žinoma, apytikris vidurkis $\bar{x}$. Pav., iš skaičių, 12,4, 12,7 ir 14,2, griežtai skaičiuodami gausime vidurkį:

$$\bar{x} = \frac{12{,}4 + 12{,}7 + 14{,}2}{3} = 13{,}1.$$

Tuos pačius skaičius apvalę nuo trupmenų gausime: 12, 13 ir 14, o jų vidurkis:

$$\bar{x} = \frac{12 + 13 + 14}{3} = 13.$$

Jei matematinis griežtumas nėra reikalingas, šis metodas su nauda gali būti vartojamas ne tik vidurkiams, bet ir kitiems statistikos skaičiams skaičiuoti.

Matematiškai griežtesnis ir metodologiškai įdomesnis antrasis vidurkių skaičiavimo būdas, kai iš visos eilutės x_i dydžių pasirenkamas kuris nors vienas dydis, tarsime a, ir vidurkis $\bar{x}$ galutinai nustatomas sulig skirtumais nuo konstantos pagal formulą:

$$\bar{x} = a + \frac{\Sigma (x_i - a)}{n} =$$

$$= a + \frac{(x_1 - a) + (x_2 - a) + \ldots + (x_i - a) + \ldots + (x_n - a)}{n} \qquad \text{(VIIIb)}$$

Sutarus x-inės eilutės dydžių x_i ir a konstantos skirtumas žymėti raide b, taigi bet kurį skirtumą

$$b_i = x_i - a,$$

minėtoji formula sutrumpėja ir gauname:

$$\bar{x} = a + \frac{\Sigma b_i}{n}.$$

Savo esmėje ta pati formula galioja ir svertiniam vidurkiui, su tuo skirtumu, kad ji papildoma x_i variantų gausumą atstovaujančiais reiškiniais y_i ir gaunama:

$$\bar{x} = a + \frac{\Sigma b_i y_i}{n} = a + \frac{\Sigma (x_i - a) y_i}{\Sigma y_i} \qquad \text{(VIIIc).}$$

$$\bar{x} = a + \frac{b_1 y_1 + b_2 y_2 + \ldots + b_i y_i + \ldots + b_n y_n}{y_1 + y_2 + \ldots + y_i + \ldots + y_n}.$$

Šis vidurkių skaičiavimo metodas pagrįstas iš aritmetikos elementoriaus žinomu faktu, kad visai vienodų dydžių aritmetinį vidurkį atstovauja bet kuris atskiras pastovus dydis. Jei visi x-inės eilutės vienetai būtų lygūs, tai jų suma Σx_i, padalinta narių skaičiumi n, duotų tą patį x_i, vadinasi, būtų $x_i = \bar{x}$.

Panašiai atsitiks, jei iš statistinės eilutės narių tarpo bet kurį x_i dydį padarysime konstanta a. Šiuo atveju tačiau reikės pirmiau surasti $x_i - a$ skirtumus b_i, ir sulig tuo, ar tų skirtumų algebrinė suma $(\Sigma x_i - a)$ išeis su teigiamu ar neigiamu ženklu, reikės, padalinus narių skaičiumi, gautąjį dalmenį prie konstantos a pridėti ar atimti. Kas dydžių su neigiamais reiškiniais vartoti nenorėtų, konstantai a gali pasirinkti patį mažiausiąjį x-inės eilutės vienetą.

Tas pats rezultatas bus gautas ir tuo atveju, kai konstantai bus paimtas didžiausias eilutės dydis. Tik šiuo atveju su neigiamais reiškiniais gautųjų skirtumų suma, padalinta narių skaičiumi, reikės nuo a konstantos atimti pagal formulą:

$$\bar{x} = a - \frac{\Sigma (a - x_i)}{n}.$$

Statistikos praktikoj konstantai paprastai parenkamas kuris nors apvalus skaičius tuo tikslu, kad lengviau duotųsi x-iniai skirtumai nustatomi. Konstantai tinka ir medianinis, t. y. pats vidurinis, dydis M. Tuo būdu:

$$\bar{x} = M + \frac{\Sigma (x_i - M)}{n} = a + \frac{\Sigma (x_i - a)}{n}.$$

Šiuo atveju konstantą a pavaduoja M. Be to, algebrinė medianinių skirtumų suma $\Sigma(x_i—M)$ parodo ir statistinės eilutės x pasiskaidymo kryptį (simetriška ar nesimetriška pati statistikos eilutė).

c) Aritmetinio vidurkio skaičiavimo pavyzdžiai.

Paprastas aritmetinis vidurkis $\bar{x}$ išskaičiuotas: 1) iš originalinių skaičių su griežtumu ligi vieneto, 2) iš apytikrių (apvalintų) skaičių su griežtumu ligi 100 ir 3) $\bar{x}$ iš skirtumų tarp konstantos a ir x-inės eilutės vienetų x_i pagal formulą VIIIb.

Lentelė 13.

MIRIMAI LIETUVOJE 1926—1935 MT.

Eilės Nr.	Metai	Mirimų skaičiai	Apvalinti mirimų skaičiai	Skirtumai iš originalinių skaičių ir konstantos (a=30000)	Skirtumai iš originalinių sk. ir vidurinio dydžio (a=x_5=37151)
		(x_i)	(x'_i)	(x_i—a)	(x_i—a=x_i—x_5)
1	1926	34 380	34 400	4 380	— 2 771
2	1927	38 897	38 900	8 897	1 746
3	1928	35 698	35 700	5 698	— 1 453
4	1929	39 669	39 700	9 669	2 518
5	1930	37 151	37 200	7 151	—
6	1931	37 478	37 500	7 478	327
7	1932	36 577	36 600	6 577	— 574
8	1933	32 749	32 700	2 749	— 4 402
9	1934	35 789	35 800	5 789	— 1 362
10	1935	34 595	34 600	4 595	— 2 556
	Iš viso	362 983= $=\Sigma x_i$	363 100= $=\Sigma x'_i$	62 983= $=\Sigma(x_i-a)$	— 8 527 $=\Sigma(x_i—x_5)$

Tuo būdu aritmetinis vidurkis iš originalinių skaičių gaunamas:

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{362\,983}{10} = 36\,298{,}3,$$

o iš tų pačių apvalintų skaičių:

$$\bar{x}' = \frac{\Sigma x'_i}{n} = \frac{363\,100}{10} = 36\,310{,}0.$$

Skirtumas, kaip matome, visai nedidelis. Pasikliaujant gi konstantomis a, vidurkiai išeina tie patys kaip ir iš originalinių skaičių:

$$\bar{x} = a + \frac{\Sigma(x_i - a)}{n} = 30\,000 + \frac{62\,983}{10} =$$

$$= 30\,000 + 6\,298{,}3 = 36\,298{,}3.$$

$$\bar{x} = a + \frac{\Sigma(x_i - a)}{n} = x_5 + \frac{\Sigma(x_i - x_5)}{n} = 37\,151 - 852{,}7 =$$

$$= 36\,298{,}3.$$

Aplamai, sulig skirtumais nuo kurio nors pastovaus dydžio (konstantos) a vidurkių skaičiavimas žymiai palengvėja, kai statistinės masės vienetai atstovaujami ne vienodai, ir yra reikalingas svertinis vidurkis. Šiuo atveju skirtumai tarp konstantos a ir statistinės eilutės variantų x_i dauginami variantų gausumo skaičiais y_i, o sandaugų algebrinė suma, padalinta statistinės eilutės narių skaičiumi, duoda dydį, kurį pridėjus (jei neigiamas — atėmus) prie pasirinktos konstantos a, gaunamas svertinis vidurkis pagal žinomą formulą:

$$\bar{x} = a + \frac{\Sigma(x_i - a)y_i}{\Sigma y_i} = a + \frac{\Sigma b_i y_i}{n}.$$

Tačiau tais atvejais, kai x-inės eilutės dydžių skirtumai (arba grupių intervalai) yra vienodi, svertinių vidurkių skaičiavimą galima dar žymiau palengvinti. Būtent, užuot dauginti gausumo skaičiais y_i statistinės eilutės originalinių dydžių x_i skirtumus nuo pasirinktos konstantos a, dauginami nuotoliniai x-inės eilutės sekos nr., kurie atžymimi teigiamais ir neigiamais pažymiais sulig tuo, kurioj konstantos a pusėj įrašomi. Šiuo atveju užuot ieškoti x-inės eilutės dydžių x_i ir konstantos a skirtumų b_i pagal formulą (skirtumas $b_i = x_i - a$) rašoma nuotoliniai eilutės dydžių numeriai $x_{\pm i}$ pradedant nuo konstantos a. Jei sutarsime šiuos x-inės eilutės nuotolinius dydžius žymėti x, tai mūsų formuloj $x_i - a$ skirtumus b_i atstovaus x ir tuo būdu formulos viršutinė dalis

$$\Sigma(x_i - a)\,y_i \text{ galės būti pakeista į } \Sigma x y_i.$$

Tačiau aritmetiniam vidurkiui gauti ši formula teks dar papildyti reiškiniu, atstovaujančiu dydžio grupių intervalą. Nes

konstanta a, kad taptų vidurkiu, yra reikalinga pataisos, kuri nusistato pagal x_i eilutės dydžių gausumą ir jų skirtumo laipsnį (intervalą) gretimose su pasirinkta konstanta a grupėse. Jei sutarsime nekintamą intervalą žymėti raide k, tai svertiniam aritmetiniam vidurkiui skaičiuoti formula VIII c

$$\bar{x} = a + \frac{\Sigma (x_i - a)y_i}{\Sigma y_i}$$

privalės būti pakeista į:

$$\bar{x} = a + \frac{k\Sigma xy}{\Sigma y_i} \qquad \text{(VIII d).}$$

Dar lengviau ir paprasčiau gaunamas aritmetinis vidurkis vadinamuoju s u m a v i m ų metodu pagal formulą:

$$\bar{x} = a + \frac{k(Sy_1)}{n} = a + \frac{k(s_1 - s_{-1})}{\Sigma y_i} \qquad \text{(VIII e)}$$

Jau iš formulos matyti, kad

$$Sy_1 = s_1 - s_{-1}.$$

O s_1 ir s_{-1} atstovauja kiekvienos statistinės eilutės x_i dydžio grupės vienetų y_i kumuliuotas, pradedant nuo kraštutinių ir baigiant konstantai a gretimuoju, skaičių sumas. Jei sutarsime a konstantai pasirinktojo x_i varianto vienetų skaičių žymėti y_i, tai s_1 bus gauta kumuliuojant visus y_i eilutės dydžius nuo apačios pradedant ligi y_1; s_{-1} gauti reikės tos pačios y_i eilutės dydžius kumuliuoti ligi y_{-1} pradedant nuo viršaus (t. y. su neigiamais reiškiniais).

Pavyzdžiui galima paimti duomenis apie darbininkų dienos uždarbio normas Lietuvoje, kaip jos buvo patiektos „Statistikos Pagrinduose". Suskirsčius tas darbo normas vienodais k intervalais po 25 ct. kiekvieną, gauname 17 grupių arba variantų x_i su nevienodo gausumo y_i skaičiais, kurių bendroji suma $\Sigma y_i = 427$ (t. y. toks buvo pranešimų skaičius). Arčiau prisižiūrėję mes randame, kad gausingiausia atstovaujama x_8 uždarbio norma, t. y. 3,25 lt., ir ši norma pakartojama 89 pranešimuose. Taigi randama, kad

$$x_8 = 3{,}25 \text{ lt., o } y_8 = 89$$

(sekantis dažniausias dydis pasikartoja tik 68 kartus, ir tai atsitinka $x_9 = 3{,}00$). Nors savaime konstanta a gali būti paimtas bet kuris st. eilutės x_i variantas, tačiau dėl visos eilės sumetimų rekomenduojama konstantai pasirinkti dažniausiąjį dydį (dominantą), kaip patikimiausią vidurkio reprezentantą, arba — vidurinį dydį. Jei sutarsime konstantai paimti dažniausiąjį dydį (šiuo atveju x_8), tai abiejų (x ir y) eilučių skaičiai, pažymint juos nuotoliniais eilutės nr. x, gaus šitokius reiškinius:

$$x_8 = a = x_0,\ x_7 = x_{-1},\ x_6 = x_{-2},\ \ldots,\ x_1 = x_{-7},\ \text{ir}$$

$$y_9 = x_1,\ x_{10} = x_2,\ \ldots,\ x_{17} = x_9.$$

Šitaip sužymėti x-inės eilutės nuo konstantos a nuotoliniai skaičiai x, ieškant a. vidurkio (X) pagal formulą VIII d, atstovaus x dydžius.

Lentelė 14.

LIETUVOS LAUKO DARBININKŲ DIENOS UŽDARBIS 1924 MT. PAVASARĮ.

Nr.	Dienos uždarbis litais (x_i)	Nuotoliniai x nuo dažniausia gautojo (y_8) uždarbio (x)	Uždarbio atsitikimų (pranešimų) skaičius (y_i)	Atskirų atsitikimų ir nuotolinių eil. skaičių sandauga (xy_i)	Uždarbio variantų ir jų atsitikimų sandauga (x_iy_i)
1	1,50	—7	2	— 14	3,00
2	1,75	—6	1	— 6	1,75
3	2,00	—5	30	—150	60,00
4	2,25	—4	16	— 64	36,00
5	2,50	—3	11	— 33	27,50
6	2,75	—2	52	—104	143,00
7	3,00	—1	68	— 68	204,00
8	3,25	0	89	—	289,25
9	3,50	1	60	60	210,00
10	3,75	2	33	66	123,75
11	4,00	3	33	99	132,00
12	4,25	4	15	60	63,75
13	4,50	5	7	35	31,50
14	4,75	6	4	24	19,00
15	5,00	7	1	7	5,00
16	5,25	8	4	32	21,00
17	5,50	9	1	9	5,50
	$\Sigma x_i = 59{,}50$		$\Sigma y_i = 427$	$\Sigma xy_i = -47$	$1376{,}00 = \Sigma x_iy_i$

Kai yra sutarta konstanta paimti x_8, taigi a = 3,25, tat vidurkis pagal formulą

$$\bar{x} = a + \frac{k(xy_i)}{\Sigma y_i} = 3{,}25 + \frac{-47 \cdot 0{,}25}{427} = 3{,}25 + \frac{-11{,}75}{427} =$$

$$= 3{,}25 - 0{,}0275\ (=)\ 3{,}22 \text{ lt.}$$

Lygiai tą patį vidurkį duoda ir formula VIIIa, kur

$$\bar{x} = \frac{\Sigma x_i y_i}{\Sigma y_i}, \text{ nes } \bar{x} = \frac{1376{,}00}{427}\ (=)\ 3{,}22$$

Visai tas pats vidurkis gaunamas ir sumavimų metodu.

Lentelė 15.

VIDUTINIO UŽDARBIO (svertinio vidurkio) SKAIČIAVIMO SCHEMA SUMAVIMŲ METODU.

Dienos uždarbis litais (x_i)	Uždarbio atsitikimų (pranešimų) skaičius (y_i)	Nulinės eilės kumuliuotos sumos (S_0) ($s_{-0} + y_0 + s_0 = Sy_0$)	Pirmosios eilės kumuliuotos sumos (S_1) ($s_1 - s_{-1} = Sy_1$)
1,50	2	2	2
1,75	1	3	5
2,00	30	33	38
2,25	16	49	87
2,50	11	60	1 47
2,75	52	112	2 59
3,00	68	$s_{-0} = 180$	$s_{-1} = 4\ 39$
3,25	89	$y_0 = 89$	—
3,50	60	$s_0 = 158$	$s_1 = 3\ 92$
3,75	33	98	2 34
4,00	33	65	1 36
4,25	15	32	71
4,50	7	17	39
4,75	4	10	22
5,00	1	6	12
5,25	4	5	6
5,50	1	1	1

Tuo būdu aritmetinis vidurkis sumavimų metodu pagal formulą VIIIe,

$$\bar{x} = a + \frac{k(Sy_1)}{n} = a + \frac{k(s_1 - s_{-1})}{\Sigma y_i},$$

duoda tą patį vidurkį (3,22). Nes kai $s_1 = 3{,}92$, $s_{-1} = 4{,}39$, o $k = 0{,}25$, $n = 427$ ir $a = 3{,}25$, tai, simbolių vieton įstatę skaičius, gauname:

$$\bar{x} = 3{,}25 + 0{,}25\frac{3{,}92 - 4{,}39}{427} = 3{,}25 + \frac{-0{,}47 \cdot 0{,}25}{427} =$$

$$= 3{,}25 - 0{,}0375 (=) 3{,}22.$$

Iš patiektosios schemos taip pat matyti, kad pirmoji, t. y. nulinė kumuliatyvinė suma

$$Sy_0 = s_{-0} + y_0 + s_1$$

yra lygi bendrajam y-inės eilutės vienetų skaičiui, nes, įstatę skaičius, gauname:

$$Sy_0 = 180 + 89 + 158 = 427.$$

Taigi $Sy_0 = N$. Ši aplinkybė, taip pat kitos sumavimų metodo ypatybės, kaip vėliau pamatysime, daro jį parankiu ir kitose statistinio analizio srityse.

Pasirenkant vieną ar kitą vidurkių skaičiavimo metodą reikia nepamiršti, kad ne tik sumavimų metodas (formula VIIIe), bet ir sutrumpintas skaičiavimo būdas pasikliaujant konstanta a ir jos nuotoliniais skaičiais x (formula VIIId) yra reikalingas vienodų intervalų (k) st. eilutės dydžiams x_i (arba vienodo dydžio skirtumų). Jeigu x_i dydžių bei jų grupių skirtumai (arba intervalai) nevienodi, tai reikia tenkintis formula VIIIc, kur

$$\bar{x} = a + \frac{\Sigma(x_i - a)y_i}{\Sigma y_i}$$

Tačiau ši formula reikalauja, kad pirmiau būtų ieškoma x_i ir konstantos a dydžių skirtumų ir tik paskui sandaugų sumos, kuri dar reikalinga padalinti, kad gauti vidurkio konstantai a reikalingą pataisą. O mes žinome, kad

$$\bar{x} = a + \frac{\Sigma(x_i - a)y_i}{\Sigma y_i} = \frac{\Sigma x_i y_i}{\Sigma y_i}.$$

Taigi svertinis vidurkis daug paprasčiau gaunamas pagal formulą VIIIa, nes ten pakanka x_i dydžius padauginti ir gautąją sumą padalinti narių skaičiumi. Dėl šios priežasties VIIIc formula išsimoka naudoti tik ypatingais atvejais (pav., jeigu x_i dydžiai yra pareikšti daugelio skaitmenų skaičiais ir konstantos a skirtumais operuojant išvengiama „astronominių" skaičių sandaugos).

d) Vidurkiai iš dydžio grupių bei klasių.

Visai atskirą problemą sudaro vidurkiai iš dydžio grupių bei klasių. Pav., statistikos šaltiniuose labai dažnai patiekiami kiekybinės klasifikacijos duomenys, bet neparodoma jokio vidurkio. Gerai, jei statistinėje lentelėje yra patiekta kiekvienos dydžio grupės vienetų skaičius ir, šalimais, tą skaičių atitinkanti kiekybinio pažymio mato vienetų bendroji suma, k. a. kiekvienos dydžio grupės ar klasės žemės ūkių bendrasis plotas, kiekvienos darbininkų kategorijos bendrasis uždarbis ar p. Tokiais atvejais iš bendrųjų skaičių lengvai gaunama ir vidurkiai. Bet gana dažnai statistikos šaltiniuose patiekiami dydžio grupių ar klasių intervalai ir vienetų skaičiai, bet nėra duomenų apie dydžio grupių apimties matą. Pav., statistikos šaltiniuose galima rasti gyventojų amžiaus grupes, bet nėra nei atskirų grupių, nei visų gyventojų bendrojo amžiaus. Patiekiami mirusiųjų, vedusiųjų, nuteistųjų ar kt. amžiaus grupių skaičiai, tačiau nėra nei atskirų grupių, nei bendrojo visų gyventojų amžiaus metų. Taip pat galima rasti darbo algos, turto ar pajamų dydžio grupes su atitinkančiais vienetų skaičiais, bet nėra bendrosios pajamų ar kt. sumos ir negalima žinoti vidurkio. Visais tokiais ar panašiais atvejais tenka ieškoti bent apytikrių vidurkių aplinkiniais keliais.

Šitie „aplinkiniai keliai" gali eiti tik per kiekybinių (dydžio) grupių ar klasių v i d u r į (medianą). Ir visais tais atvejais, kai dydžio grupės yra sudarytos statistiškai matematiniu metodu, t. y. kai grupių intervalai aplamai yra vienodi ir kraštutinių grupių intervalai taip pat aprėžti ar bent apytikriai aprėžtini — labai didelės paklaidos įvykti negali. Daug sunkiau yra tikslingai apibrėžti kraštutinių, t. y. griežtai neaprėžtų, dydžio grupių bei klasių vidurinį dydį. Tokiais atvejais

apie vidurinį dydį galima tik spėti. Ir jeigu tokios grupės sk. vienetai labai išsiskiria (labai stambūs ar labai smulkūs, lyginant su visa mase), tai gali nemaža klaida įvykti.

Lentelė 16.

LIETUVOS ŽEMĖS ŪKIAI DYDŽIO KLASĖMIS.

Nr.	Dydžio klasės ($k_{x,i}$)	Ūkių skaičius (y_i)	Plotas ha ($y_i\ k_{x,i}$)	Dydžio klasių ūkių vidurkiai $\left(\bar{x}=\frac{k_{x,i}}{y_i}\right)$	Dydžio klasių viduriniai dydžiai $\left(\ddot{x}_i=\frac{k_1+k_2}{2}\right)$	Plotas (ha) iš vidurinių dydžių ($\ddot{x}$) ($\ddot{x}_i y_i$)
0	1	2	3	4	5	6
1	1— 2	13 797	18 862	1,37	1,5	20 695,5
2	2— 5	39 666	135 214	3,41	3,5	138 831,0
3	5— 8	45 882	295 307	6,44	6,5	298 233,0
4	8— 10	32 355	287 590	8,89	9,0	291 195,0
5	10— 12	28 060	304 612	10,86	11,0	308 660,0
6	12— 15	31 512	419 123	13,30	13,5	425 412,0
7	15— 20	33 236	569 921	17,15	17,5	581 630,0
8	20— 30	34 197	826 593	24,17	25,0	854 925,0
9	30— 50	20 597	762 894	37,04	40,0	823 880,0
10	50—100	6 476	425 025	65,63	75,0	485 700,0
11	100—150	906	108 154	119,38	125,0	113 250,0
12	150—200	379	64 169	169,31	175,0	66 325,0
13	200 ir d.	317	101 050	318,77	245,0	77 665,0
	Iš viso	287 380	4 318 514	$\bar{x}=15{,}03$		4 486 401,5

Skaičiavimo pavyzdžiu galima paimti Lietuvos žemės ūkių 1930 m. surašymo duomenis, kaip jie yra paskelbti statistikos šaltiniuose. Čia mes turime statistikos lentelę, kur dydžio klasės patiektos skirtingais intervalais ir kur pati stambiausiųjų dydžių klasė (ūkiai nuo 200 ha) su 317 vienetų (ūkių) palikta neaprėžta. Ši lentelė patogi tuo, kad statistikos šaltiniuose galima rasti duomenų gautajam apytikriam vidurkiui patikrinti.

Statistiškai metodologiniu atžvilgiu patikimiausias aritmetinis vidurkis, kai žinomi tik dydžio grupių skaitomieji vienetai, turėtų išeiti, kiekvienos dydžio grupės variantu x_i paėmus

atitinkamos grupės intervalo vidurkį, t. y. jo vidurinį dydį. Jei sutarsime dydžio klasės intervalo žemutinę ribą žymėti k_1, o aukštutinę — k_2, tai statistinėj eilutėj bet kurios grupės x_i dydžių spėjamas reprezentantas (vidurinis dydis) $\ddot{x}$ bus gaunamas pagal formulą:

$$\ddot{x} = \frac{k_1 + k_2}{2} = k_1 + \frac{k_2 - k_1}{2} \qquad \text{(IX).}$$

Ši formula, tačiau, neįgalina aprėžti $\ddot{x}$, kai k_1 ar k_2 palieka neaprėžti. Pav., mūsų paimtame pavyzdyje $k_{13,1} = 200$ ha, bet visai nėra žinių apie $k_{13,2}$. Grynai teoretiškai imant, $k_{13,2}$ gali siekti ir 201 ir 20 000 ha, ar net daugiau.

Tokiu atveju k_2 aprėžiant reikia atsižvelgti anksčiau buvusių dydžio klasių intervalo kitimų proporcijos. Jeigu priešpaskutinės dydžio klasės yra aprėžtos vienodais intervalais, o jų vienetų skaičiai mažėja, tat yra pagrindo tarti, kad ir šios dydžio klasės (k_n) vidurinis dydis $\ddot{x}$ bus maždaug tokiame pat atstume nuo anksčiau buvusios (k_{n-1}) klasės, kaip ši, lyginant ją su priešpaskutine (k_{n-2}).

Kadangi mūsų paimtame pavyzdyje $\ddot{x}_{12}$, lyginant su $\ddot{x}_{11}$, klasės lyginys

$$1_{12} = \frac{\ddot{x}_{12}}{\ddot{x}_{11}} = \frac{175}{125} = 1{,}40$$

tai yra pagrindo teigti, kad ir 1_{13} perdaug neatsitolins nuo 1,40 = $= \frac{\ddot{x}_{13}}{175}$, t. y. galės būti apie 245 ha, nes jeigu $\frac{\ddot{x}_{13}}{175} = 1{,}40$, tai $\ddot{x}_{13} = 175 \cdot 1{,}40 = 245$.

Šiuo pagrindu sudarytais dydžio klasių intervalais visų žemės ūkių aritmetinis vidurkis išeina 15,61 ha, nes

$$\bar{x}'' = \frac{\Sigma \ddot{x}'_i y_i}{\Sigma y_i} =$$

$$= \frac{4\,486\,401{,}5}{287\,380{,}0} = 15{,}61 \text{ ha.}$$

Tikrasis aritmetinis vidurkis išeina 15,03 ha, nes bendrasis visų ūkių plotas yra ne 4 486 401,5 ha, o tik 4 318 514 ha. Taigi

paklaida (skirtumas tarp $\bar{x}$ ir $\bar{x}''$) sudaro 0,58 ha. (15,61—15,03= = 0,58) arba nepilnus 4% $[\frac{0{,}58 \cdot 100}{15{,}03} (=) 3{,}86]$. Tačiau iš pridedamos kontrolinės lentelės 16 mes galime įsitikinti, kad šią paklaidą nulėmė net ne tiek paskutinios dydžio klasės interpoliuotas vidurkis, kiek kitų (aprėžtais intervalais) klasių interpoliuotieji vidurkiai ($\bar{x}$). Tikrasis paskutinės dydžio klasės vidurkis net išeina didesnis (beveik 319 ha). Taigi mūsų paimtasis (245 ha) spėjamąjį $\bar{x}'$ dar sumažina, t. y. pataiso. Matyti, kad dydžio klasių dauguma buvo sudaryta neatsižvelgiant grupuojamų dydžių pasiskaidymo apie vidurinį dydį (medianą).

e) Medianas.

Medianas, arba medianinis dydis, yra pats vidurinis bet kurios skaičių eilutės dydis.

Norint medianą surasti, yra reikalinga duotąją skaičių eilutę išrikiuoti pagal jos narių svorį, didumą arba, aplamai, pagal vyresniškumą. Taip išrikiuotoj eilutėj patsai vidurinis dydis ir bus medianinis dydis. Jei duotojoj st. eilutėj atskirų dydžių (narių) skaičius lyginis, tai medianui imamas abiejų vidurinių dydžių vidurkis. Pav., skaičių eilutėj, kur atskiri dydžiai vienas po kito seka kaip 5, 6, 7, *8*, 9, 10, 11, medianą atstovaus pačiam vidury esantis skaičius *8*. Skaičių gi eilutėj vienu nariu pailgintoj (5, 6, 7, *8*, *9*, 10, 11, 12) medianinį dydį gausime ieškoti abiejų vidurinių dydžių tarpe, nes čia bus st. eilutės vidurys. Šiuo atveju M = 8, 5, nes

$$\frac{8+9}{2} = 8{,}5.$$

Medianinio dydžio eilutės nr. nustatyti galioja formula:

$$M_n = \frac{N+1}{2} \qquad (X).$$

Šioj formuloj N atstovauja bendrąjį st. eilutės narių skaičių. Formula palengvina surasti medianą ypač tais atvejais, kai originaliniai dydžiai jau sugrupuoti į dydžio grupes arba klases. Pavyzdžiui galime imti tuos pačius Lietuvos žemės ūkių surašymo duomenis, kur ūkių skaičius, t. y. N = 287 380. Tuo būdu vidurinio ūkio M dydis (ha skaičius) pareis nuo

$$M_n = \frac{287\,380 + 1}{2} = 143\,690{,}5 \text{ ūkio.}$$

O kadangi pusės ūkio tikrenybėje nėra, tai M teks interpoliuoti, t. y. paimti pačių vidurinių ūkių (143 690-jo ir 143 691-jo) aritmetinį vidurkį.

Tuo būdu mediano (M) ieškojimas klasifikuotoj dydžio grupių eilutėj (arba lentelėj) eina dviem, arba net trimis, etapais:

1) suradimas vidurinio st. eilutės nr.,
2) suradimas tos dydžio grupės arba klasės y_i, kurios M_n priklauso,
3) nustatymas surastojo vidurinio (M_n) dydžio charakteristikos (jo mato vienetų skaičių).

Pirmasis etapas atliekamas pagal jau žinomą formulą:

$$M_n = \frac{N + 1}{2}.$$

Medianinės dydžio klasės tenka ieškoti kumuliuojant dydžio grupes (y_i) ligi tol, kol komuliuotųjų grupių bendroji suma Sy_i bus nemažesnė kaip M_n. Jeigu atsitiktų, kad $Sy_i = M_n$, tai paskiausiai priimtosios y_i grupės intervalo k_i aukštutinė riba k_2 atitiktų medianinį dydį M (arba $M_{\bar{x}}$). Bet jeigu $Sy_i > M_n$ (paskiausiai priimtoji grupė duotų skaičių, didesnį už M_n), tai medianinis dydis tektų nustatyti iš y_i grupių atitinkančio intervalo k_i proporcionaliai tam, kiek reikalinga iš šios grupės paimti skaitomųjų vienetų, kad pasiekti patį vidurinį st. eilutės narį (M_n).

Jei y_i klasės intervalą atžymėsime k_i, jo kraštutines mato ribas k_1 ir k_2 ir paskutinę kumuliuotųjų grupių vienetų sumą Sy_i, tai x-inės eilutės medianinį dydį M galėsime nustatyti pagal formulą:

$$M = k_{i,2} + \frac{s_i k_i}{y_i} = \quad = k_{i\,2} + \frac{M_n - Sy_i}{y_i}(k_{i,2} - k_{i,1}) \qquad \text{(XI).}$$

Formula rodo, kad $M_{\tilde{x}}$ gali atitikti M_n dydį dviem atvejais: a) kai $M_n = Sy_i$ ir b) kai $k_1 = k_2$ (kai M_n eilutės narį atitinkanti dydžio klasė susideda iš vienodo dydžio vienetų).

Pavyzdžiui paimkime tuos pačius žemės ūkių surašymo duomenis. Jau matėme, kad $M_n = 143\,690{,}5$. Lentelėje 16 randame, kad ūkių dydžio klasės turi:

	Ūkiai y_i	Kumuliuotos sumos $Sy_i = y_{i+1}$	Intervalai $k_{i,2} - k_{i,1} = k_i$	
y_1	13 797	13 797	1 — 2 ha	1
y_2	39 666	53 463	2 — 5 „	3
y_3	45 882	99 345	5 — 8 „	3
y_4	32 355	131 700	8 — 10 „	2
y_5	**28 060**	**159 760**	**10 — 12 „**	2
y_6	31 512		12 — 15 „	3

ir t. t.

Kaip matome, šeštosios grupės vienetai medianui ieškoti nebereikalingi, nes jau $Sy_5 = 159\,760$, t. y. duoda skaičių, kuris yra didesnis už M_n. Iš kitos, tačiau, pusės negalima tenkintis ir pirmųjų keturių grupių kumuliuota suma, nes Sy_4, apimanti visus dydžius ligi 10 ha, tepasiekia tik 131 700 eil. nario; taigi ligi M_n (143 690) dar trūksta daugiau kaip 11 tūkstančių. Penktojoj gi klasėj (k_5) su intervalo ribomis tarp 10—12 ha, tačiau, yra iš viso net 28 060 vienetų ($y_5 = 28\,060$). Kai visa tai žinome, tat ir medianą pagal XI formulą gauname:

$$M = k_{5,2} + \frac{M_n - Sy_5}{y_5}(k_{5,2} - k_{5,1})$$

$$= 12 + \frac{143\,690{,}5 - 159\,760}{28\,060}(12 - 10)$$

$$= 12 + \frac{-16\,069{,}5 \cdot 2}{28\,060}$$

$$(=)\ 12 - 1{,}14\ (=)\ 10{,}86 \text{ ha.}$$

Jei dabar paimsime skaičius iš lentelės 14 (apie darbo algą), tai gausime šitokius duomenis:

$N = 427$, taigi

$$M_n = \frac{N + 1}{2} = \frac{427 + 1}{2} = 214$$

$y_1 = 2$, taigi $Sy_i = 2$ ir $x_1 = 1{,}50$
$y_2 = 1$, „ „ $= 3$ „ „ $= 1{,}75$
$y_3 = 30$, „ „ $= 33$ „ „ $= 2{,}00$
$y_4 = 16$, „ „ $= 49$ „ „ $= 2{,}25$
$y_5 = 11$, „ „ $= 60$ „ „ $= 2{,}50$
$y_6 = 52$, „ „ $= 112$ „ „ $= 2{,}75$
$y_7 = 68$, „ „ $= 180$ „ „ $= 3{,}00$
$y_8 = \mathbf{89}$, „ „ $= \mathbf{269}$ „ „ $= \mathbf{3{,}25}$
$y_9 = 60$

Mūsų formula XI šiuo atveju duoda:

$$M = k_{8,2} + \frac{M_n - Sy_8}{y_8}(k_{8,2} - k_{8,1})$$

$$= 3{,}25 + \frac{214 - 269}{89}\ (3{,}25 - 3{,}25)$$

$$= 3{,}25 + \frac{-55 \cdot 0}{89}$$

$$= 3{,}25 + 0$$

$$= 3{,}25.$$

Matome, kad formula XI tinka bet kokiam atvejui.

Savaime suprantama, kad galimi ir kiti šios formulos varijantai. Nes kai

$$k_i = k_{i,2} - k_{i,1},\ \text{o}$$

$$M = k_{i,2} + \frac{(M_n - Sv_i)\,k_i}{y_i},$$

tai formulą galima ir taip parašyti:

$$M = k_{i,2} - \frac{(Sy_i - M_n)\,k_i}{y_i} \qquad \text{(XIa)}$$

Tie patys rezultatai bus gauti ir pagal šią formulą:

$$M = k_{i,1} + \frac{(M_n - Sy_{i-1})\,k_i}{y_i} \qquad \text{(XIb)}$$

Pastarojoj formuloj M pagrindą sudaro žemutinė intervalo riba, o skirtumas s_i gaunamas ir medianinio skaičiaus M_n ir kumuliuotųjų dydžio grupių bendrosios sumos, kuri gaunama be tos dydžio grupės, kurioj randamas medianinis dydis.

f) Dominantas.

Dominantas, anglų ir iš dalies rusų dažnai vadinamas modus (Mode), vokiečių yra vadinamas dažniausiuoju arba gausingiausiuoju dydžiu (häufigste, dichteste Wert), t. y. tokiu dydžiu, kuris dažniausia sutinkamas arba dažniausia pasikartoja. Iš to seka, kad jeigu kurioj nors statistikos eilutėj kiekvienas dydis tepasikartoja tik vieną kartą, tai dominanto, t. y. dominuojančio dydžio nėra. Taip pat nebus dominanto statistinėj eilutėj, kur visi dydžiai pasikartoja vienodai. Iš kitos, tačiau, pusės, jei kuris nors dydis pasikartoja daugiau, negu bet kuris kitas atskirai paimtas dydis, tai tuo pačiu jis paliktų dominantu, — nors tai būtų ir patsai mažiausias arba patsai didžiausias, — t. y. kuris nors kraštutinis eilutės dydis.

Ši visai logiška išvada iš daug kur vartojamo dominuojančio dydžio apibrėžimo ne visai sutampa su ta dominanto sąvoka, kuri į statistiką pateko iš matematikų žodyno. Kai L. March ar kiti jį pasekusieji statistikai kalba apie dominantą, arba Yule, Bowley ar kt. anglai rašo apie modusą (Mode), tai, paprastai, jie turi galvoje dydį, kuris gali būti gausingiausias ir kuris mažiau ar daugiau sutampa su st. dydžių maksimaline tikimybe normaliame statistinės eilutės dydžių pasiskaidyme apie medianą bei aritmetinį vidurkį.

Iš tikrųjų, moduso arba dominanto sąvoka yra susieta su binomo koeficientais bei tikimybėmis. Šia prasme dominantas tampa aritmetinio vidurkio bei mediano konkurentu. O tokiu konkurentu negali būti nė vienas kraštutinis statistinės eilutės dydis. Dėl to ir nevadinsime dominanto dažniausiuoju nei gausingiausiuoju dydžiu, nors etimologiškai toks, vokiečių pavyzdžiu, sulietuvinimas ir atrodytų tikslingas. Kalbėsime apie dominantą tik matematinės statistikos prasme ir vadinsime juo tik tokį dažniausia pasikartojantį dydį, kuris gali būti randamas mediano artimoj kaimynystėj. Tokiais atvejais dominantas (D) duodasi apytikriai apibrėžiamas pagal formulą

$$D (=) \bar{x} - 3 (\bar{x} - M) \qquad \text{(XII)}.$$

Sąmoningai šią formulą rašau, priešingai daugeliui kitų statistikų, tik su apytikrės lygybės ženklu (=). Nes dominan-

tas socialekonominio gyvenimo reiškiniuose, paprastai, nėra matematiškai griežtame santykyje nei su medianu, nei su aritmetiniu vidurkiu. Teisybė, jei $\bar{x} = M$, tai $D = \bar{x} - 3(\bar{x} - M)$, $\bar{x} - M = 0$, ir tuo būdu ir $D = M = \bar{x}$, — vadinasi, aritmetinis vidurkis, medianas ir dominantas sutampa. Dominanto statistinis apibrėžimas turi prasmės tik tuo atveju, kai dydžiai yra susikaupę ne griežtai simetriškai, o kurioj nors vienoj pusėj nepertoli nuo mediano vienodų dydžių yra daugiau. Jei statistinėj eilutėj daugiau yra smulkesniųjų dydžių, tai aritmetinis vidurkis $\bar{x}$ bus už *M* didesnis, — ir mažesnis, jei gausingesni stambesnieji dydžiai. Formula: $D (=) \bar{x} - 3(\bar{x} - M)$ rodo, kad tarpas tarp dominanto ir aritmetinio vidurkio, prie apynormalės dispersijos, yra apie tris kartus didesnis, kaip tarpas tarp aritmetinio vidurkio ir mediano. Iš to seka, kad jeigu yra žinomas aritmetinis vidurkis ($\bar{x}$) ir dominantas (D), tai duodasi surandamas ir medianas (M) pagal formulą:

$$M (=) \bar{x} + \frac{D - \bar{x}}{3} \qquad \text{(XIII).}$$

Vadinasi, jei statistinė eilutė bent kiek (bet ne perdaug) asimetriška, tai, dominanto ir aritmetinio vidurkio skirtumo trečdalį pridėję prie aritmetinio vidurkio, gausime medianą. Likusieji gi du trečdaliai to paties $(D - \bar{x})$ skirtumo sudaro jau nuotolį tarp mediano (M) ir dominanto (D). O iš to seka, kad, žinant dominantą (D) ir medianą (M), galime sužinoti ir apytikrį aritmetinį vidurkį pagal formulą:

$$\bar{x} (=) M - \frac{D - M}{2} \qquad \text{(XIV).}$$

Pavyzdžiui paimkime mūsų aukščiau turėtus žemės ūkių surašymo duomenis, kur aritmetinis vidurkis $\bar{x} = 15{,}03$, o medianas $M = 10{,}86$. Šiose aplinkybėse turėtų būti:

$$D = \bar{x} - 3(\bar{x} - M) = 15{,}03 - 3(15{,}03 - 10{,}86) =$$
$$= 15{,}03 - 3 \cdot 4{,}17 = 2{,}52 \text{ ha.}$$

Dabar patikrinkime.

Jei $D = 2{,}52$, o $\bar{x} = 15{,}03$, tai medianas (M) turėtų būti:

$$M = \bar{x} + \frac{D - \bar{x}}{3} = 15{,}03 + \frac{2{,}52 - 15{,}03}{3} =$$

$$= 15{,}03 - \frac{12{,}51}{3} = 15{,}03 - 4{,}17 = 10{,}86.$$

Taip pat ir $\bar{x}$ iš D ir M:

$$\bar{x} = M - \frac{D - M}{2} = 10{,}86 - \frac{2{,}52 - 10{,}86}{2} =$$

$$= 10{,}86 + \frac{8{,}34}{2} = 10{,}86 + 4{,}17 =$$

$$= 15{,}03 \text{ ha.}$$

Mūsų paimtas pavyzdys, tačiau, taip pat parodo, kad dominantui surasti formula bent socialekonominėj statistikoj turi aprėžtas pritaikymo ribas. Iš mediano ir aritmetinio vidurkio nustatę dominantinį dydį, mes randame, kad šis dydis išeina apie keturis kartus mažesnis už medianą ir beveik šešis kartus mažesnis už aritmetinį vidurkį (santykio lyginiai kaip 1 : 4 : 6). Toks didelis skirtumas savaime rodo labai nesimetrišką statistinės eilutės dydžių pasiskaidymą. Kai mažiausias ūkis mūsų duomenyse turi 1 ha, medianas beveik 11 ha, o aritmetinis vidurkis 15 ha, tai dominantas iš 2,52 ha rodo, jog smulkiųjų ūkių svoris turi būti žymiai didesnis, ne kaip rodo aritmetinis vidurkis. Berods, iš 16 lentelės patiektųjų duomenų matyti, kad dydžio klasių intervalai nėra vienodi ir dėl to griežtai empiriškai dominanto negalima nustatyti. Vis dėlto jau tie duomenys duoda pagrindo abejoti, kad 2,52 ha ūkių būtų daugiausia. Tiesa, 2-je dydžio klasėje (tarp 2—5 ha), kur pagal formulą išeina dominantas, tėra 39 666 ūkiai, kai sekančioj klasėj (5—8 ha), yra parodyta 45 882, t. y. daugiau negu bet kurioj kitoj klasėj. Tačiau ir šios klasės intervalo 1 ha išeina 45 882 : 3 = 15 294 ūkiai, kai sekanti dydžio klasė (tarp 8—10 ha) su 2 ha intervalu turi 32 355 ūkius ir intervaliniam ha tenka 32 355 : 2 = = 16 177,5 ūkiai.

Lietuvos žemės ūkio surašymo šaltiniuose galima rasti trijų apskričių (Mažeikių, Šakių ir Ukmergės) duomenis, suskirstytus į dydžio grupes vienodais 1 ha intervalais. Šie duomenys labai gerai tinka dominanto ir kitų vidurkių metodologinėms

problemoms išaiškinti (žiūr. Visuotinas Lietuvos Žemės Ūkio surašymas 1930 m. I, psl. 28). Šiose apskrityse būta 38 696 ūkių 1 — 794 ha ribose su 539 063 ha bendro ploto. Dydžio grupių su 1 ha intervalu iš viso būta 207, o gausingiausioji grupė tarp 6—7 ha buvo 2 919 ūkiai, ir jų vidurkis siekė 6,44 ha $\left(\frac{18\,810 \text{ ha}}{2\,919} (=) 6{,}44 \text{ ha}\right)$. Jei žiūrėti ne tik ūkių skaičiaus, bet ir jų ploto bendroje ūkių masėje, tai dominuoja dydžio grupė tarp 8—9 ha, kur ūkių buvo 2 653, t. y. mažiau ne grupėj tarp 6—7 ha; bet ploto šioj grupėj buvo 22 338 ha, t. y. daugiau.

Tuo tarpu aritmetinis visų ūkių vidurkis išeina 13,93 ha $\left(\frac{539\,063}{38\,696} (=) 13{,}93 \text{ ha}\right)$. Medianas tenka 19 348,5 ūkiui ($M_n = \frac{N+1}{2} = \frac{38\,696+1}{2} = 19\,348{,}5$) ir tai išeina dydžio grupei tarp 9—10 ha. Griežtai skaičiuojant pagal formulą gaunama:

$$M = k_{i,2} + \frac{(M_n - Sy_i)\, k_i}{y_i}$$

$$= 10 + \frac{19\,348{,}5 - 20\,128}{2\,121} \cdot (10—9) =$$

$$= 10 + \frac{-779{,}5}{2\,121} (=) 10—0{,}35 =$$

$$(=) 9{,}65 \text{ ha.}$$

Pamėginkime dabar empiriškai gautaisiais dydžiais patikrinti $\bar{x}$, M, D dydžių savitarpio santykį pagal duotąją formulą. Jeigu visais atvejais būtų $D = \bar{x} - 3(\bar{x} - M)$, ir visa, kas toliau iš tos formulos seka, tai mūsų pavyzdyje dominantas turėtų išeiti:

$$D (=) \bar{x} - 3(\bar{x} - M) = 13{,}93 - 3(13{,}93 - 9{,}65) =$$

$$= 13{,}93 - 3 \cdot 4{,}28 = 13{,}93 - 12{,}84 (=) 1{,}09 \text{ ha!}$$

O tai jau neįmanomas dydis — net jei nežinotume, kad empiriškai mūsų nustatytas dominantas D (=) 6,44 ha.

Taip pat jeigu išeiti iš mūsų empiriškai surastųjų D ir $\bar{x}$ dydžių, tai medianui gausime ne empiriškai surastąjį M = 9,65 ha, o 11,43 ha. Nes jeigu D = 6,44, o $\bar{x}$ = 13,93, tai M pagal formulą (žiūr. pusl. 181):

Lentelė 17a.

ŽIEMINIŲ KVIEČIŲ DERLIAI 1932—34 M.

Nr.	Apskritys	Derliai 100 kg iš 1 ha (x_i)	Plotas 1000 ha (y_i)	Bendras derlius 1000 kvintalų ($x_i y_i$)
1.	Telšiai	10,5	2,2	23,10
2.	Utena	11,0	4,6	50,60
3.	Seinai	11,2	2,0	22,40
4.	Trakai	11,3	2,4	27,12
5.	Zarasai	11,4	1,3	14,82
6.	Ukmergė	11,5	6,8	78,20
7.	Kaunas	11,6	10,8	125,28
8.	Alytus	11,9	7,8	92,82
9.	Kretinga	11,9	2,6	30,94
10.	Šakiai	12,6	8,5	107,10
11.	Panevėžys	12,7	15,2	193,04
12.	Marijampolė	13,0	10,6	137,80
13.	Tauragė	13,3	4,5	59,85
14.	Rokiškis	13,3	4,0	53,20
15.	Kėdainiai	13,3	14,4	191,52
16.	Raseiniai	13,4	10,7	143,38
17.	Mažeikiai	13,6	7,1	96,56
18.	Biržai	13,6	12,8	174,08
19.	Šiauliai	13,7	24,4	334,28
20.	Klaipėda	14,7	1,2	17,64
21.	Vilkaviškis	15,1	8,2	123,82
22.	Šilutė	15,7	0,1	1,57
23.	Pagėgiai	18,2	0,4	7,28
	Visoj Lietuvoj $\bar{x}$ (=)	12,95	Σy_i = 162,6	$\Sigma x_i y_i$ = 2 106,40

$$M\,(=)\,\bar{x} + \frac{D - \bar{x}}{3} = 13{,}93 + \frac{5{,}44 - 13{,}93}{3} =$$

$$= 13{,}93 - \frac{7{,}49}{3}\,(=)\,11{,}43 \text{ ha.}$$

Skirtumas tarp empiriškai surastojo ir hipotetinio M šiuo atveju nėra toks didelis, kaip anksčiau dėl D, bet vis dėlto gana žymus ir toli gražu neatitinka tikrenybę. O iš to jau seka, kad statistikos praktikoj tais atvejais, kai dispersija anormali, dominantas taip pat yra sunkiai apibrėžiamas ir dar sunkiau surandamas dydis, kaip ir gyvenime vadinamieji „dažniausieji" atvejai arba „vidutiniai" dydžiai. Šios kategorijos priklauso ir Quetelet „vidutinis žmogus".

Tačiau tais atvejais, kai statistinės eilutės dispersija artima normaliai, formula gali būti su nauda vartojama. Pavyzdžiui galime paimti Lietuvos kviečių derliaus duomenis 1932—34 m.

Žieminių kviečių 1932—1934 m. vidutinio derliaus skirtumai atskiruose Lietuvos apskrityse svyravo tarp 10,5—18,2 kvintalų iš 1 ha. Svertinis vidurkis visai Lietuvai išeina:

$$\bar{x} = \frac{1}{N}\Sigma x_i y_i = \frac{2\,106{,}40}{162{,}6}\,(=)\,12{,}95 \text{ q iš 1 ha.}$$

Kadangi šiuo atveju N = 162,6 (=) 163, tai $M_n = \frac{162{,}6}{2} =$ = 81,8. Kumuliuodami visų y_i dydžius nuo x_1 ligi 14 (Rokiškio apskr.) mes gauname $Sy_i = 79{,}3$ — bet tai yra mažiau negu M_n (81,8). Sekantis derlius sutinkamas jau 14,4 t. ha. Taigi medianinio dydžio eil. nr. (M_n) tenka y_{15} grupėj, kur derliai iš 1 ha sudarė 13,3 q = x_{15}. O kadangi ir šios grupės visi dydžiai vienodi (13,3) tai ir $M = x_{15} = 13{,}3$ ha.

Dominantą gi gauname pagal XII formulą:

$$D\,(=)\,\bar{x} - 3\,(\bar{x} - M) = 12{,}95 - 3\,(12{,}95 - 13{,}30) =$$
$$= 12{,}95 + 3 \cdot 0{,}35 = 12{,}95 + 1{,}05 =$$
$$= 14{,}0 \text{ q iš 1 ha.}$$

Iš lentelės matyti, kad didžiausias atskirai paimtas plotas Šiaulių apskr. su 24,41 ha davė 13,7 q iš 1 ha. Taigi šiuo at-

veju hipotetinis ir empiriškai nustatytas D gana artimi dydžiai. Tuo būdu formula ir socialekonominėj statistikoj tam tikrose sąlygose yra pritaikoma.

Matematinė statistika turi metodų ir priemonių ištirti ir nustatyti, ar reikiamos sąlygos yra ar jų nėra. Bet apie tai teks vėliau kalbėti. Tuo tarpu dar susipažinsime su viena statistikos vidurkių kategorija — su vad. matematiniais vidurkiais.

g) Matematiniai vidurkiai[1].

Matematiniais vidurkiais vadinami vidurkiai, kurie skaičiuojami pasikliaujant kurio nors aukštesniojo laipsnio lygčių sistema. Tokių lygčių bei jų sistemų, o tuo pačiu ir vad. matematinių vidurkių, matematikai turi nemaža. Statistikos praktikoj, tačiau, dažniau yra vartojami tik geometriniai vidurkiai, nors savaime šie vidurkiai nėra nei pranašesni, nei statistikai prasmingesni už bet kuriuos kitus vad. matematinius vidurkius. Čia mes išnagrinėsime, be geometrinių, dar ir vad. harmoninius, kvadratinius ir antiharmoninius vidurkius, kurie statistikos vadovėliuose dažniau sutinkami.

Geometriniai vidurkiai bet kuriai x-inei statistikos eilutei skaičiuojami pagal formulą:

$$\bar{x}_g = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \ldots x_n}, \text{ arba}$$

$$\log\bar{x}_g = \frac{\log x_1 + \log x_2 + \log x_3 + \ldots\ldots + \log x_n}{n} =$$

$$= \log\bar{x}_g = \frac{\Sigma(\log x_i)}{n} = \frac{1}{n}\Sigma(\log x_i) \qquad \text{(XV).}$$

Formula rodo, kad vartojant logaritmus geometrinio vidurkio skaičiavimo technika iš esmės nesiskiria nuo aritmetinio vidurkio skaičiavimo. Taip pat ta pati formula pritaikoma ir sver-

[1] Matematinių vidurkių prasmingumo klausimai statistikos vadovėliuose, paprastai, nenagrinėjami, o pasitenkinama tik nurodymu, kaip jie santykiuoja su aritmetiniu vidurkiu. Šiuo atžvilgiu esminės išimties nesudaro net naujausieji matematinės statistikos autorių darbai (plg. *H. L. Rietz-Bauer,* Handbuch der mathematischen Statistik, pp. 36/7, *O. N. Anderson,* Einführung in die mathematische Statistik, 145/9, 180 ir kt., *Czuber-Burkhardt,* Die statistischen Forschungsmethoden, 96—107, ir kt.).

tiniams vidurkiams. Šiuo atveju, kaip ir svertiniams aritmetiniams vidurkiams, galioja:

$$\log\bar{x}_g = \frac{y_1\log x_1 + y_2\log x_2 + y_3\log x_3 + \ldots + y_n \log x_n}{y_1 + y_2 + y_3 + \ldots + y_n} = \frac{\Sigma y_i \log x_i}{\Sigma y_i} \qquad \text{(XVa).}$$

Savo prigimtimi geometrinis vidurkis yra visuomet mažesnis už aritmetinį vidurkį ($\bar{x}_g < \bar{x}$), o už geometrinį vidurkį dar mažesnis vad. h a r m o n i n i s vidurkis.

H a r m o n i n i s vidurkis skaičiuojamas pagal formulą:

$$\bar{x}_h = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \ldots + \frac{1}{x_n}} = \frac{n}{\Sigma \frac{1}{x_i}} \qquad \text{(XVI).}$$

Skaičiavimo technika, kaip rodo formula, čia dar paprastesnė: x-inės eilutės narių skaičius n dalomas atverstųjų (reciprokinių) x_i dydžių bendrąja suma.

Jei $\bar{x}_g$ ir $\bar{x}_h$ už $\bar{x}_a$ yra mažesnis ir santykiuoja kaip $\bar{x}_h < \bar{x}_g < \bar{x}_a$, tai kvadratinis ir antiharmoninis vidurkiai ($\bar{x}_q$ ir $\bar{x}_{ah}$) yra už aritmetinį vidurkį didesni ir santykiuoja kaip $\bar{x}_a < \bar{x}_q < \bar{x}_{ah}$.

K v a d r a t i n i s vidurkis skaičiuojamas pagal formulą:

$$\bar{x}_q = \sqrt{\frac{x^2_1 + x^2_2 + \ldots + x^2_n}{n}} = \sqrt{\frac{1}{n}\ \Sigma x^2_i} \qquad \text{(XVII).}$$

A n t i h a r m o n i n i a m vidurkiui galioja formula:

$$\bar{x}_{ah} = \frac{x^2_1 + x^2_2 + \ldots + x^2_n}{x_1 + x_2 + \ldots + x_n} = \frac{\Sigma x^2_i}{\Sigma x_i} \qquad \text{(XVIII).}$$

Pati skaičiavimo technika, kaip matome, ir čia nėra perdaug komplikuota. Aplamai, jei paimti bet kurią skaičių eilutę ir surasti aritmetinį, geometrinį, harmoninį, kvadratinį ir antiharmoninį vidurkius, tai gausime šitokią atskirų eilutės vidurkių dydžių rikutę $\bar{x}_h < \bar{x}_g < \bar{x}_a < \bar{x}_q < \bar{x}_{ah}$.

Pavyzdžiui galime paimti bet kuriuos skaičius ir patikrinti. Štai skaičiai 2, 5 ir 8.

Aritmetinis vidurkis $\bar{x}_a = 5$, nes: $\frac{2+5+8}{3} = \frac{15}{3} = 5.$

Geometrinis vidurkis ($\bar{x}_g$) išeis 4,31, nes:

$$\bar{x}_g = \sqrt[3]{2.5.8} = \sqrt[3]{80} (=) 4{,}31;$$

$$\log\bar{x}_g = \frac{0{,}3010 + 0{,}6990 + 0{,}9031}{3} = \frac{1{,}9031}{3} =$$

(=) 0,6344; o iš $\lg\bar{x}_g = 0{,}6344$ gaunama $\bar{x}_g$ (=) 4,31.

Harmoninis vidurkis pagal atitinkamą formulą išeina:

$$\bar{x}_h = \frac{n}{\Sigma\frac{1}{x_i}} = \frac{3}{\frac{1}{2}+\frac{1}{5}+\frac{1}{8}} = \frac{3}{0{,}500+0{,}200+0{,}125} = \frac{3}{0{,}825}$$

(=) 3,64.

Kvadratinis vidurkis pagal formulą:

$$\bar{x}_q = \sqrt{\frac{1}{n}\Sigma x_i^2} = \sqrt{\frac{2^2+5^2+8^2}{3}} = \sqrt{\frac{4+25+64}{3}} =$$

$$= \sqrt{31} \ (=)5{,}67$$

Antiharmoninis vidurkis tiems patiems skaičiams gaunamas:

$$\bar{x}_{ah} = \frac{\Sigma x_i^2}{\Sigma x_i} = \frac{2^2+5^2+8^2}{2+5+8} = \frac{93}{15} = 6{,}2.$$

Surašę visus penkis vidurkius į eilutę ir aritmetinį vidurkį prilyginę 100, kai n = 3, o $x_1 = 2$, $x_2 = 5$ ir $x_3 = 8$, gauname šitokio santykiavimo lyginių lentelę:

	$\bar{x}_i$	$\frac{\bar{x}_i \cdot 100}{\bar{x}_a}$	$\frac{(\bar{x}_i - \bar{x}_a)\,100}{\bar{x}_a}$
$\bar{x}_h$ =	3,64	72,8	— 27,2
$\bar{x}_g$ =	4,31	86,2	— 13,8
$\bar{x}_a$ =	5,00	100,0	0,0
$\bar{x}_q$ =	5,67	113,4	13,4
$\bar{x}_{ah}$ =	6,20	124,0	24,0

Kaip matome, čia apibrėžtieji vad. matematiniai vidurkiai yra tam tikrame ir, palyginti, gana griežtame santykyje su

aritmetiniu vidurkiu, būtent: geometrinis vidurkis už aritmetinį vidurkį beveik tiek pat mažesnis, kiek, procentais skaitant, kvadratinis vidurkis už aritmetinį vidurkį didesnis. Taip pat gana artimoj skirtumų proporcijoj yra harmoninis ir antiharmoninis vidurkiai. Šita aplinkybė pravartu gerai įsidėmėti: nes ji rodo, kad vidurkiniai dydžiai tiems patiems skaičiams pareina nuo pasirinkto vidurkių skaičiavimo metodo. O juo skaičių eilutės dydžių skirtumai didesni, tuo daugiau skirsis ir patys vidurkiniai dydžiai.

h) Vidurkių statistiniai uždaviniai ir jų tikslingas parinkimas.

Ideališkas statistikos vidurkis turi būti charakteringas statistikos skaičių eilutės (stebimos masės arba jos reiškinių) reprezentantas. Taigi atstovauti statistikos skaičių eilutę (statistikos metodais stebimą masę arba reiškinį) ir ją charakterizuoti yra pagrindinis statistikos vidurkių uždavinys. O kurios rūšies vidurkis (dominantas ar medianas, aritmetinis ar kuris nors matematinis) kurioms sąlygoms yra tinkamesnis arba tikslingesnis, tai jau statistikos teorijos bei metodologijos pareiga parodyti bei paaiškinti.

Aplamai, jei tik gerai įsižiūrėsime matematinių vidurkių formulas, tai lengvai pastebėsime, kad tose formulose yra reiškinių arba komponentų, kurie savo turiniu primena svertinių aritmetikos vidurkių reiškinius bei komponentus. Imkime, pav., kvadratinio vidurkio formulą:

$$\bar{x}_q = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \ldots + x_n^2}{n}}.$$

Ši formula pirštu prikišamai parodo, kad bet kuris skaičių eilutės dydis x_i pirmiau turi būti tam tikru koeficientu, lygiai kaip ir svertiniams vidurkiams, padaugintas, paskui visi x-iniai dydžiai susumuoti ir tik po to dalybos keliu vidurkio ieškoma. Kiekvienas gi x_i dydis dauginamas skirtingu koeficientu, be to, didesnieji dydžiai, keliant laipsniu, šioj sistemoj dauginami didesniais, o mažesnieji — mažesniais skaičiais. Dėl to neturi stebinti ir ta aplinkybė, kad šiuo metodu iš tų pačių x-inės eilutės dydžių vidurkis gaunamas didesnis už paprastąjį aritmetinį vidurkį. Iš tikrųjų, kvadratinio vidurkio for-

mulos pašaknio tiekybės skaitiklį $(x_1^2 + x_2^2 + x_3^2 + \ldots + x_n^2)$ mes turime teisę ir taip parašyti:

$$x_1x_1 + x_2x_2 + x_3x_3 + \ldots + x_n\,x_n.$$

Šitaip gi mes rašėme ir svertinio vidurkio formulos skaitiklį, su tuo skirtumu, kad daugiklius x_i tenai žymėjome y_i. Ir jeigu kvadratiniam vidurkiui gauti skaitiklį $(\Sigma x_i x_i)$ dalome ne Σx_i, o tik n, ir paskui ištraukiame šaknį, tai tuo būdu pasiekiame tik to, kad gauname kiekybę, kuri yra tam tikru procentu mažesnė už svertinį vidurkį. Nes jeigu paimsime mūsų anksčiau turėtus skaičius — 2, 5 ir 8 — tai svertinį vidurkį pagal formulą $\bar{x} = \frac{\Sigma x_i y_i}{\Sigma y_i}$ gausime:

n	x_i	y_i	$x_1\,y_1$
1	2	2	4
2	5	5	25
3	8	8	64
3	$15 = \Sigma x_1$	$15 = \Sigma y_i$	$93 = \Sigma x_1 y_1$

$$\bar{x} = \frac{\Sigma x_i\,y_i}{\Sigma y_1} = \frac{93}{15} = 6{,}2.$$

Pasižiūrėję į atskirų vidurkių lentelę psl. 184, mes rasime, kad tiems patiems dydžiams šiais svorio koeficientais gautasis svertinis vidurkis visai atitinka vad. antiharmoninį vidurkį. Patikrinę gi vad. antiharmoninio vidurkio formulą mes įsitikinsime, kad toji formula, iš tikrųjų, atitinka svertinio vidurkio formulos visas sąlygas. Nes iš esmės nebus jokio skirtumo, ar mes rašysime:

$$\bar{x} = \frac{\Sigma x_i^2}{\Sigma x_i}, \quad \bar{x} = \frac{\Sigma x_i x_i}{\Sigma x_i} \quad \text{ar} \quad \bar{x} = \frac{\Sigma x_i y_i}{\Sigma y_i}.$$

Ir jei tik visais atvejais $x_i = y_i$, tai nebus jokio skirtumo ir vidurkio dydžiuose.

Tos pačios sąlygos nulemia ir geometrinius bei vad. harmoninius vidurkius, su tuo skirtumu, kad šiais atvejais svorio

koeficientai didesni išeina smulkesniems dydžiams. Pav., imkime vėl tuos pačius skaičius 2, 5 ir 8 ir išskaičiuokime svertinius vidurkius dviem varijantais:

A			B		
x_i	y_i	$x_i y_i$	x_i	y_i	$x_i y_i$
2	10	20	2	5	10
5	7	35	5	7	35
8	5	40	8	10	80
	$22 = \Sigma y_i$	$95 = \Sigma x_i y_i$		$22 = \Sigma y_i$	$125 = \Sigma x_i y_i$

Tuo būdu tie patys x-inės eilutės dydžiai, sukeitus svorio koeficientus, duos svertinio vidurkio variantus:

$$\bar{x}_a = \frac{\Sigma x_i y_i}{\Sigma y_i} = \frac{95}{22} (=) 4{,}32 \text{ ir}$$

$$\bar{x}_b = \frac{125}{22} (=) 5{,}68.$$

Kaip matome, parinkus tam tikrus svorio koeficientus tiems patiems dydžiams, gaunami svertiniai vidurkiai, kurie visai artimi geometriniam bei kvadratiniam vidurkiui.

Ši aplinkybė yra labai svarbi ir visais atvejais įsidėmėtina. Vadinasi, jei imti bet kurios kilmės statistinę eilutę, kur yra daugiau ar mažiau skirtingų dydžių (variantų) x_i, tai tiems dydžiams išskaičiuotieji — matematinis ir paprastas aritmetinis — vidurkiai skirsis taip pat, kaip ir skirtųsi paprastas aritmetinis vidurkis nuo svertinio vidurkio, — prie sąlygos, jei matematiniam vidurkiui skaičiuoti būtų parinkta statistinės eilutės dydžių pasiskaidymą (dispersiją) atitinkanti formula. Ką tik patiektieji svertinių vidurkių variantai A ir B iš dalies apibrėžia ir matematinių vidurkių pritaikymo sąlygas. Jei matematinį vidurkį būtų norima kiek galint priartinti svertiniam vidurkiui ar dominantui, kai yra duomenų ar rimto pagrindo teigti, kad smulkesnieji dydžiai gausingiau atstovaujami, tenka imti geometrinį arba harmoninį vidurkį; o jeigu masėje yra daugiau stambesniųjų dydžių, jos charakterį geriau reprezentuos kvadratinis arba antiharmoninis vidurkis.

Aplamai, tačiau, imant, tenka turėti galvoje, kad bet kurios statistinės eilutės „ideališkiausias" ir matematiškai patikimiausias buvo ir pasiliks aritmetinis, resp. svertinis a. vidurkis[2]. Nes tik aritmetinis vidurkis visais be išimties atvejais atitinka pagrindinę griežtai matematinio skaičiavimo sąlygą, kuri reikalauja, kad bet kurios skaičių eilutės reprezentantų suma atitiktų pagrindinių dydžių visumą. Jeigu $\bar{x}$ turi atstovauti $x_1, x_2, x_3, \ldots, x_n$ eilutės visus narius ir $\bar{x} = \frac{1}{n} \Sigma x_i$, tai vidurkis $\bar{x}$, padaugintas eilutės narių skaičiumi, turi būti lygus tos pačios x_i eilutės visumai. Vadinasi, $\Sigma x_i = \bar{x}n$. Bet jeigu tai pačiai x_i dydžių eilutei, šalia aritmetinio vidurkio $\bar{x}$, dar kuriuo kitu metodu išskaičiuosime vidurkį (medianą, geometrinį, kvadratinį ar kt.) ir jeigu gausis $M < \bar{x}$ (arba $\bar{x} < M$), tai bus ir $Mn < \bar{x}n$ (arba $\bar{x}n < Mn$), o ne tapatybės, kokias sudaro $\Sigma x_i = \bar{x}n$.

Vadinasi, visais tais atvejais, kai statistinės eilutės skirtingųjų dydžių pasiskaidymo (svorio) koeficientai yra žinomi ir svertinių vidurkių skaičiavimas perdidelių technikos sunkenybių nesudaro, svertiniams a. vidurkiams reikia atiduoti pirmenybę.

Tas pats klausimas, tačiau, jau visai kita plotme tenka svarstyti ir spręsti tais atvejais, kai statistinės eilutės x_i dydžiai tėra apytikriai arba net nevienodu matu matuojami. O tai sutinkama gana dažnai, ypač socialinio ir ekonominio gyvenimo reiškinių statistikoj, kai duomenys statistikos šaltiniuose patiekiami jau nevienodo dydžio grupėse ar net labai skirtingais intervalais dydžio klasėse. Net tais atvejais, kai iš patiektųjų duomenų šiose aplinkybėse galima svertinį aritmetinį vidurkį surasti, jisai apie visą masę ar reiškinį vienas nedaug ką pasakys, jei nebus žinoma bent pagrindinės dydžių dispersijos ypatybės. O tų dispersijos ypatybių iš vieno a. vi-

---

[2] Kitas reikalas, pav., fizikoj arba mechanikoj, kur traukos, svorio, greičio arba jėgos kitimai visuomet vyksta dėsningai. Jeigu tuos kitimų variantus pažymėti skaičiais, tai gaunama skaičių eilutė iš dėsningai pasiskaidžiusių dydžių; o žinant šių dydžių pasiskaidymo dėsningumą, galima ir jų vidurkį, kuris atitiktų kitimų arba skirtumų dėsningumus, vien matematikos (be statistinio stebėjimo) priemonėmis apibrėžti.

durkio, jei nėra žinomi x_i dydžių svorio koeficientai, bent kiek patikimiau net ir nustatyti negalima.

Tačiau apie dispersijos pobūdį ir pagrindinę kryptį jau galima apytikriai spręsti iš aritmetinio vidurkio santykiavimo su medianu. Apytikris gi medianas (M) galima surasti iš statistinių dydžių grupių arba klasių ir su skirtingais intervalais (žiūr. pusl. 173/4).

Turint svertinį a. vidurkį (m) ir medianą (M) pagal žinomą (psl. 176) formulą XII galima apytikriai nustatyti ir dominantas (D). Iš šių gi trijų dydžių jau galima susidaryti ir apytikrį vaizdą apie visos statistinės eilutės pobūdį net ir tuo atveju, jeigu kitų pagalbinių priemonių, k. a. nuo vidurkio arba mediano nukrypimų, arba Galtono metodo dispersijos pobūdžiui nustatyti panaudota nebūtų.

Vadinamieji matematiniai vidurkiai tais atvejais, kai svertinis a. vidurkis yra žinomas arba kai statistikos duomenys yra gauti ištisinio stebėjimo metodu, nedaug ką naujo gali pasakyti. Juo labiau, kad matematiniai vidurkiai šiose sąlygose (kai labai dideli skaitomųjų vienetų kiekiai sugrupuoti skirtingų intervalų dydžio klasėse) iš viso būtų gana problematiškos vertės, be to, ir jų skaičiavimas būtų gana sunkus ir daug darbo reikalingas. Kitas tačiau reikalas, kai turimoji statistikos medžiaga yra gauta reprezentaciniu metodu, arba kai statistinės eilutės dydžių (nors ir labai skirtingų), yra, palyginti, nedaug, ir tik po vieną kiekvieno. Šiose aplinkybėse jau matematiniai vidurkiai gali statistinės prasmės turėti. Imkime pavyzdžiui skaičių eilutę iš penkių x_i dydžių, kurie sudaro šitokią eilutę:

$$x_1 = 10,\ x_2 = 12,\ x_3 = 13,\ x_4 = 17,\ x_5 = 18.$$

Šitokiai eilutei lengva nustatyti, kad M = 13 (patsai vidurinis, t. y. x_3, dydis), o aritmetinis vidurkis

$$\bar{x} = 14, \text{ nes } \frac{10 + 12 + 13 + 17 + 18}{5} = \frac{70}{5} = 14.$$

Tarkime, kad šie penki x_i dydžiai yra atsitiktinai paimti iš bet kurios masės, — vadinasi, yra atsitiktiniai masės repre-

zentantai[3]. Jau a. vidurkio ($\bar{x}$) ir mediano (M) nesutapimas duoda pagrindo įtarti, kad pačioje masėje šių dydžių pasiskaidymas nėra griežtai normalus. Paprastai tokiais atvejais būna, arba galima rasti, dar ir kitų simptomų arba duomenų, kurie rodo, jog dydžių pasiskaidymas nėra nei ideališkai normalus, nei griežtai simetriškas. Vienas iš tokių simptomų, tai skirtumai nuo mediano ($x_i - M$).

Jeigu $\Sigma (x_i - M)$ gaunama ne 0, o bet kuris pozityvinis ar negatyvinis skaičius, tai jau simptomas, kad ir pačioje masėje, greičiausia, stambesnieji ar smulkesnieji dydžiai gausingesni arba, kaip sakoma, svaresni. Mūsų atveju kai M = 13, randame:

		$x_i - M$
x_1	10	—3
x_2	12	—1
x_3	**13**	0
x_4	17	4
x_5	18	5
		$\Sigma(x_i - M) = 5$

Vadinasi, šiuo atveju $x_i - M$ pozityviniai skirtumai yra svaresni. Vadinasi, dominantas (D), pagal mums jau žinomą taisyklę, turėtų būti mažesnis ir už medianą (M), o patsai skirtumas tarp dominanto ir mediano turėtų būti dvigubai didesnis už skirtumą tarp mediano ir aritmetinio vidurkio. Pritaikę atitinkamą formulą gauname:

$$D (=) \bar{x} - 3 (\bar{x} - M) =$$
$$= 14 - 3 (14 - 13) = 14 - 3 . 1 =$$
$$= 11.$$

Štai šioj vietoj statistikas ir turi rimtai pagalvoti, ką daryti su reprezentaciniu metodu gautaisiais skaičiais. Jei tikrai do-

---

[3] Nors savaime penketo reprezentantų ir, palyginti, nedidelei masei atstovauti nepakanka, bet principiniam klausimui išaiškinti ir penkių skaičių pakanka, — juo labiau, kad skaičiavimo rezultatai bus tie patys, jei kiekvieną x_i dydį vienodu svorio koeficientu padauginsime, vadinasi, ne penkis, o 50 ar 5000 x_i vienetų paimsime.

minantas, kad ir ne tiek daug, yra mažesnis už medianą, tai yra pagrindo spėti, kad nedaugelio dydžių reprezentantams paprastąjį aritmetinį vidurkį gali būti nulėmę tie du atsitiktinai patekę stambesnieji dydžiai, kurie pačioje masėje, gali būti, visai nėra gausingi. Vadinasi, yra pagrindo spėti, kad šiuo atveju paprastasis a. vidurkis iš nedaugelio reprezentantų nėra ideališkas visos masės reprezentantas, ir kad svertinis a. vidurkis, greičiausia, būtų mažesnis. O jeigu taip, tai čia gali praversti jau matematinis vidurkis, šiuo atveju geometrinis arba harmoninis: nes jie duos dydžius, kurie bus artimesni medianui ir dominantui, — vadinasi, ir svertiniam a. vidurkiui, negu paprastas a. vidurkis[4]. Jeigu būtų M už $\bar{x}$ didesnis, tai, žinoma, geriau tiktų kvadratinis arba antiharmoninis vidurkis (atsižvelgiant į tai, kaip dideli M ir $\bar{x}$ skirtumai ir, be to, kuriam dydžiui — medianui ar dominantui — norima vidurkį daugiau priartinti).

Aplamai, tenka įsidėmėti, kad matematiniai vidurkiai bet kuriai statistikos eilutei charakterizuoti negali būti imami savivalingai, o privalo atitikti statistinės eilutės ypatybes. Šito nepaisant, matematinis vidurkis gali dar daugiau nutolti nuo tipiškų statistikos eilutės masės dydžių ir nuo (dominanto), kaip paprastasis aritmetinis vidurkis, kai x-inės eilutės dydžiai yra nevienodai pasiskaidę.

Apskritai imant, duotąją skaičių eilutę norint charakterizuoti vidurkiais, tenka vadovautis šiais samprotavimais bei dėsniais:

1. Jei skaičių eilutė yra pakankamai didelė, o atskirų dydžių skirtumai nežymūs, galima tenkintis paprastu aritmetiniu vidurkiu, kuris lengviausia ir greičiausia surandamas. Tokiose aplinkybėse aritmetinis vidurkis negalės daug skirtis nei nuo svertinio vidurkio, nei nuo dominanto, nei nuo mediano.

Aplamai, aritmetinis vidurkis atitiks svertinį vidurkį, ir medianą ir tais atvejais, kai atskirų dydžių skirtumai bus didesni, bet patys dydžiai su dominantu pryšakyje bus simetriškai pasiskaidę apie medianą.

---

[4] Pritaikę atitinkamą formulą rasime, kad, pav., $\lg \bar{x} = \frac{1}{5}$ (lg 10 + lg 12 + lg 13 + lg 17 + lg 18) = 1,13575; atlogaritmavę $\bar{x}$ (=) 13,67.

2. Svertinis aritmetinis vidurkis skaičiuotinas visais tais atvejais, kai dydžių skirtumai yra dideli, o jų dažnumas (svorio koeficientai) yra žinomas.

3. Matematiniai vidurkiai skaičiuotini tais atvejais, kai dydžių skirtumai yra žymesni, o svorio koeficientai svertiniams vidurkiams išskaičiuoti nežinomi (arba jų nustatymas bei sulig jais vidurkių išskaičiavimas perdaug darbo reikalautų).

4) Matematinius vidurkius skaičiuojant tenka galvoje turėti, kad:

a) geometriniai (ir harmoniniai) vidurkiai geriau charakterizuos asimetriškai pasiskaidžiusią statistikos eilutę su gausingiau atstovaujamais smulkesniais dydžiais, o

b) kvadratiniai (ir antiharmoniniai) vidurkiai bus tinkamesni asimetriškai pasiskaidžiusiai statistikos eilutei, kur gausingiau atstovaujami stambesnieji dydžiai.

c) Gausingumo rodikliais čia gali būti medianas ir paprastas aritmetinis vidurkis. Jeigu aritmetinis vidurkis už medianą mažesnis, tai bus žymė, kad medianas yra daugiau stambesniųjų dydžių įtakoje, vadinasi, dydžių susikaupimas didesnis dešinėje M pusėje; čia dominantą geriau atitiks kvadratinis arba antiharmoninis vidurkis. Ir atvirkščiai: jei už a. vidurkį medianas mažesnis, statistiškai prasmingesnis bus geometrinis arba harmoninis vidurkis.

Tačiau vienais vidurkiniais dydžiais statistikos eilučių, t. y. jų atstovaujamų masių arba masinių reiškinių, pasiskaidymo ypatybių (vad. statistikos eilučių dispersijos) charakterio bent kiek tiksliau apibrėžti negalima.

6. Statistikos eilučių dydžių pasiskaidymas bei dispersija.

a) St. dispersijos sąvoka ir rūšys (tipai).

Bet kurios masės arba ją atstovaujančios statistinės eilutės charakterį nulemia atskirų x_i dydžių skirtumai ir jų pasiskaidymas arba vad. dispersija. O kadangi statistinė masė ir ją atstovaujanti eilutė paprastai, arba pirmiausia, charakterizuojama bei matuojama vidurkiais, tai ir jos variantų (x_i dydžių) skirtumai bei jų dispersiją taip pat tikslingiausia lyginti bei matuoti santykyje su vidurkiais. Sulig tuo, ar duo-

tajai statistikos eilutei x_i dydžių skirtumai nuo vidurkio yra didesni ar mažesni ir, be to, sulig tuo, kiek x_i dydžių yra susikaupę tam tikroj atokumoj nuo vidurkio, gali būti sprendžiama ir apie vidurkio mažesnį ar didesnį tinkamumą būti visos statistinės masės, arba statistinės eilutės, reprezentantu.

Apie statistinės eilutės dydžių pasiskaidymą šiek tiek galima spręsti ir iš pačių vidurkių. Jei dominantas ir aritmetinis vidurkis sutampa (vienas antrą atitinka), tai bus jau pažymys, kad statistinės eilutės dydžiai yra pasiskaidę simetriškai apie medianą. Tačiau gali atsitikti, kad statistinės eilutės dydžiai bus simetriškai apie medianą pasiskaidę ir net medianas atitiks aritmetinį vidurkį, ir vis dėlto dominantas, t. y. gausingiausias dydis, bus gana atokiai nuo abiejų. O taip atsitiks visais tais atvejais, kai bet kurioje masėje gausingesni susikaupimai bus dviejose vietose (bus maždaug po vienodai pačių smulkiųjų ir pačių stambiųjų dydžių), o viduryje, t. y. apie medianą, dydžių bus visai nedaug.

Pavyzdžiui paimkime skaičius 10, 15 ir 20. Aritmetinis vidurkis ir medianas visai sutaps, ar mes imsime:

$$\bar{x}_1 = \frac{10 + 15 + 20}{3} = \frac{45}{3} = 15, \text{ ar}$$

$$\bar{x}_2 = \frac{10 \cdot 3 + 15 \cdot 10 + 20 \cdot 3}{3+10+3} = \frac{240}{16} = 15, \text{ ar}$$

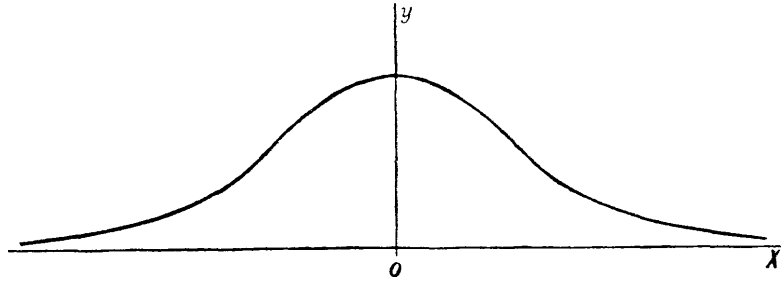
$$\bar{x}_3 = \frac{10 \cdot 10 + 15 \cdot 3 + 20 \cdot 10}{10 + 3 + 10} = \frac{345}{23} = 15.$$

Tačiau matome, kad $\bar{x}_1$ yra gauta iš trijų x_i variantų, kurie st. eilutėj atstovaujami vienodai, vadinasi, $\bar{x}_1 = M_1$. Bet čia negalima nustatyti dominanto, nes kiekvienas dydis tik vieną kartą pasikartoja. Sekantis $\bar{x}_2$ gautas iš st. eilutės, kur vidurinis dydis $x_2 = 15$ yra daugiau ne trigubai gausingesnis už kraštutinius ($y_2 = 10$, kai y_1 ir $y_3 = 3$). Taigi, čia aritmetinis vidurkis sutampa ne tik su medianu, bet ir su dominantu ($\bar{x}_2 = M_2 = D_2$). Trečiuoju atveju patsai vidurinis dydis (15) pasikartoja tik 3 kartus, o kraštutiniai (10 ir 20) — po 10 kartų. Vadinasi, į dominantą turi lygias teises dydžiai 10 ir 20, bet aritmetinis vidurkis ($\bar{x}$) ir medianas (M) palieka kaip buvę ($\bar{x}_3 = M_3$). Iš to seka, kad patsai vienas atskirų vidurkių rū-

šių santykio žinojimas dar nepakankamas, kad būtų galima spręsti apie statistinės masės pasiskaidymo pobūdį.

Aplamai, bet kurią statistinę masę (st. eilutę) pagal jos vienetų pasiskaidymo pobūdį galima priskirti vienam iš trijų pagrindinių pasiskaidymo arba dispersijos tipų.

Vieną dispersijos tipą sudaro tokios statistinės masės, kur kiekybinėj klasifikacijoj vidutinių dydžių dauguma yra simetriškai pasiskaidžiusi apie medianą, o smulkesnieji ir stambesnieji dydžiai sutinkami juo rečiau, kuo daugiau jie skiriasi nuo mediano. Šios rūšies statistinę masę vaizduojant grafiškai ortogonalių koordinatų sistemoje gaunama charakteringa kreivė, kurios aukščiausias taškas ordinatoj y atitinka horizontalinės koordinatos ašį x_0, ir nuo šio taško į abi puses kreivė leidžiasi taisyklingai žemyn.

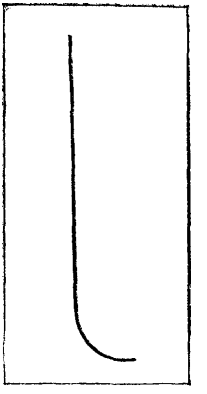


Šios rūšies ir pavidalo statistinės eilutės dydžių dispersija (pasiskaidymas) vadinamas n o r m a l i a, ir jos gryną arba klasišką tipą atstovauja binomo koeficientų dispersija.

Taip pat beveik ideališkai normalia dispersija pasižymi ypatingai visi vadinamieji natūriniai reiškiniai fizikoj, biologijoj ir kt. Tačiau tokia pat normaliai artutine dispersija pasižymi ir daugelis socialinio ir net ekonominio gyvenimo reiškinių ypač demografinėj, agronominėj, kainų ir kt. statistikoj.

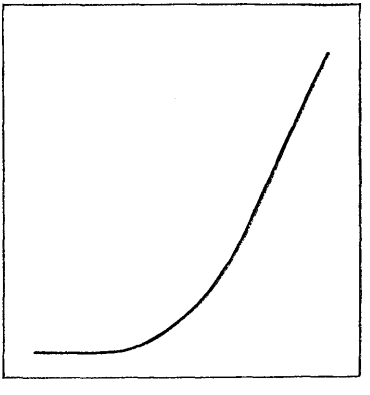
Iš kitos, tačiau, pusės ūkio gyvenime yra nemaža masių ir masinių reiškinių, kur vad. normalios dispersijos žymių nematyti. Bet ir čia dažniausiai galima išskirti du statistinės masės dydžių pasiskaidymo tipai: a) kai statistinės masės dydžių dauguma yra susikaupusi apie kurį nors vieną iš pačių kraštutinių dydžių (smulkiausiąjį ar stambiausiąjį dydį) ir b) kai statistinėj masėj iš viso yra ne vienas, o du susikaupimo dydžių sektoriai, iš kurių vienas koncentruojasi smulkiausiųjų

dydžių sektoriuj, o antras — stambiausiųjų dydžių sektoriuj.

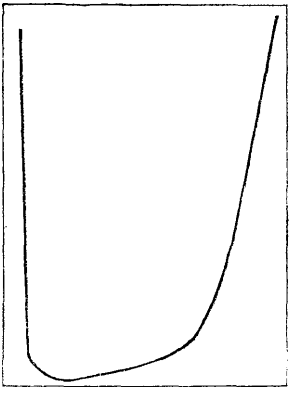


Pirmojo (a) tipo charakteringi pavyzdžiai sutinkami pajamų arba darbo algos statistikoj, kur smulkiausiųjų pajamų gavėjų visuomet yra daugiausia, ir atvirkščiai: juo didesnės kieno pajamos, tuo mažiau randama panašaus didumo pajamų gavėjų. Koordinatų sistemoj nubrėžus kreivę gaunama charakteringoji vad. P a r e t o kreivė, panaši į atvirkščią J (patiektoji kreivė nubrėžta pagal Lietuvos gyventojų mirtingumą nuo tymų 1933 m. amžiaus grupėmis).

Antrojo tipo reiškinių charakteringa dispersija sutinkama mirimų statistikoj, nes, kaip žinoma, dažniausias mirtingumas pasireiškia jauniausiojo ir seniausiojo amžiaus grupėse. Čia kreivė išeina panaši į U. Dėl to kai kas ir vadina Pareto kreivės vaizduojamą dispersiją *J* dispersija, o antrąjį anormalios dispersijos tipą — *U* dispersija.



Lietuvos gyventojų mirtingumas amžiaus grupėmis nuo reumatizmo ir podagros (1933 m.).



Lietuvos gyventojų mirtingumas amžiaus grupėmis 1933 m.

Bet kurios statistinės masės dispersijos charakterį geriausia gali parodyti kiekybinė klasifikacija griežtai vienodais, bet neperdideliais intervalais. Nes jeigu paimti perdidelius intervalus, tai charakteringosios statistinės masės ypatybės gali išnykti. Tačiau statistikos šaltiniuose neperdideliais ir griežtai vienodais intervalais sugrupuotų dydžio grupių retai sutinka-

ma. Dažniausia sutinkamos dydžio klasės su nevienodais intervalais. O iš taip sudarytų dydžio klasių masės pasiskaidymo visų charakteringų ypatybių, reikalingų griežtesniam duomenų analiziui, galima nepastebėti. Be to, visais tais atvejais, kai duomenys yra gauti reprezentaciniu metodu, statistinės eilutės dydžių pasiskaidymą ir visos eilutės dispersijos charakteringas ypatybes reikalinga kiek galint griežčiau ir nuodugniau ištirti, kad būtų galima iš turimos reprezentacinės medžiagos spręsti apie visą statistinę masę.

Šiems ir kitiems statistinio analizio uždaviniams, visų pirma, yra reikalinga ištirti ir objektyviškai nustatyti, ar statistinę masę atstovaujanti eilutė yra normali, ar nenormali, ir jeigu nenormali, tai kurio būtent anormalios dispersijos tipo priklauso. Nes jeigu statistinės eilutės dydžių dispersija (pasiskaidymas apie vidurkį) yra bent artima normaliai, tai tolimesniam st. analiziui ir event. išvadoms galioja tam tikri matematiškai pagrįsti statistinio analizio dėsniai, kuriais galima pasikliauti bei pasitikėti. Jei dispersija nenormali, tai tolimesnis st. analizis reikalingas tokių priemonių, kurios atitinka anormalios dispersijos tipo (J ar U) charakterį. Iš to seka, kad pirmasis statistinio analizio uždavinys — ištirti ir nustatyti statistinės eilutės dispersijos pobūdį ir patikrinti vidurkio, kaip st. eilutės reprezentanto, patikimumą bei tinkamumą visos st. masės reprezentantu būti.

Iš statistinių dydžių bei reiškinių šiam reikalui, be įvairaus tipo vidurkių, pirmaeilės reikšmės turi statistinės eilutės dydžių pasiskaidymo pobūdį vaizduojantis a s i m e t r i j o s koeficientas. Asimetrijos koeficientas rodo, simetriškai ar nesimetriškai st. eilutės dydžiai yra pasiskaidę apie medianą. Be to, ne mažiau svarbus yra d i s p e r s i j o s koeficientas, kuris turi parodyti, normalus ar nenormalus yra dydžių susikaupimas apie vidurkį ir medianą. Be to, dispersijos koeficientas gali parodyti, ligi kiek, arba ligi kokio laipsnio, yra tinkamas ar netinkamas patsai vidurkis statistinei masei (statistikos eilutei) reprezentuoti.

Patsai paprasčiausias ir techniškai lengviausias būdas statistinės eilutės dispersijos pobūdžiui bendraisiais bruožais pažinti — tai pasikliaut medianu ir nuo jo nukrypimais. Jau psl.

177—9 buvo parodyta, kad st. eilutės atskirų dydžių ir mediano (M) skirtumų algebrinė suma $\Sigma\ (x_i - M)$ pakankamai aiškiai parodo, kurioj mediano pusėj duotomis sąlygomis tenka ieškoti dažniausia sutinkamo dydžio (dominanto). Tačiau tos rūšies skirtumų suma dar maža ką pasako apie pačią dispersiją, nes $\Sigma(x_i - M)$ (skirtumų algebrinė suma) gali ta pati susidaryti tiek iš dydžių, esančių mediano artimiausioj kaimynystėj, tiek iš pačių atokiausių (tolimiausių) dydžių. Vadinasi, ta pati skirtumų suma gali būti ir tuo atveju, kai dydžių dauguma yra susikaupusi arti mediano, reiškia, ir dispersija yra mažiau ar daugiau normali, — tiek tuo atveju, kai medianui artimų dydžių yra visai maža, o gausingiausi dydžių susikaupimai susikoncentravę prie kraštutinių, arba atokiausiųjų, dydžių.

Šios rūšies trūkumo neturi vad. k v a r t i l i n i a i dydžiai.

b) Kvartiliai ir kvartiliniai koeficientai.

Kvartiliniai statistinės eilutės dydžiai gaunami visai tuo pačiu būdu, kaip ir medianas. Vadinasi, pirmiausia st. eilutės dydžiai turi būti išrikiuoti sulig vyresniškumu nuo mažiausiojo ligi didžiausiojo, o paskui jau ieškomi kvartiliniai dydžiai, kurie, paprastai, žymimi raide *Q*. Tokių Q yra du — pirmasis Q_1, kuris visuomet randamas pačiam viduryj tarp mediano ir mažiausiojo dydžio, ir antrasis Q_2, kurį atstovauja pats vidurinis dydis tarp mediano ir kito kraštutinio — paties didžiojo — vieneto.

Jei statistikos eilutės x_i dydžiai yra po vieną sulig vyresniškumu išrikiuoti ir kiekvienas yra atžymėtas eilutės nr. kaip

$$x_1, x_2, x_3, \ldots, x_i, \ldots, x_N,$$

tai kvartiliams Q, kaip ir medianui M, galioja formula:

$$Q_{n,1} = \frac{N+1}{4} = Q_{x,1} \qquad \text{(XIX a).}$$

$$Q_{n,2} = \frac{3\,(N+1)}{4} = Q_{x,2} \qquad \text{(XIX b).}$$

Vadinasi, kaip M dydį tokiais atvejais atstovauja x_n, gautas pagal formulą $M_n = \frac{N+1}{2} = M = x_n$, taip pat ir $Q = x_n$

gaunamas pagal formulą XIX. Tačiau jeigu statistikos eilutės dydžiai yra sugrupuoti į dydžio klases, tai, kaip ir medianui panašiose sąlygose, patiektosios formulos teparodys tik eilės nr., o patys kvartiliniai dydžiai dar reikės surasti papildomu skaičiavimu panašiai kaip ir medianą.

Atitinkamą XIX formulą pakeitę gauname:

$$Q = k_{i,2} + \frac{(Q_n - Sy_i)k_i}{y_i} = k_{i,2} - \frac{(Sy_i - Q_n)k_i}{y_i} \qquad (XX).$$

Pavyzdžiui galime paimti 16 tabelės žemės ūkio surašymo duomenis, kur dydžio klasių (k) yra 13, o ūkių (N) 287 380. Kvartilinių dydžių eilės nr. Q_n pagal atitinkamas formulas gausime:

$$Q_{n,1} = \frac{1}{4}(N + 1) = \frac{287\,380 + 1}{4} = 71\,845{,}25.$$

$$Q_{n,2} = \frac{1}{4}(N + 1)\,3 = \frac{(287\,380 + 1) \cdot 3}{4} = 71\,845{,}25 \cdot 3 =$$

$$= 215\,535{,}75.$$

Kumuliuodami atskirų dydžio klasių y_i ūkių sumas randame, kad Sy_i dėl Q_1, atitinka

$Sy_3 = y_1 + y_2 + y_3 = 99\,345$, kai $Q_{1,n} = 71\,845{,}25$. Vadinasi, pirmasis kvartilis (Q_1) tenka trečiajai dydžio klasei (taigi $y_i = y_3$), kurios intervalas k_i yra 5 — 8 ha ribose. Taigi $k_i = 3$ (nes $k_{3,1} = 5$ ir $k_{3,2} = 8$, taigi $k_i = k_{3,2} - k_{3,1} = 8 - 5 = 3$).

Tuo būdu Q_1, pakeitę XX formulos simbolius skaičiais gauname:

$$Q_1 = k_{3,2} - \frac{(Sy_3 - Q_{1,n})\,k_3}{y_3}$$

$$= 8 - \frac{99\,345 - 71\,845{,}25}{45\,882}(8-5) =$$

$$= 8 - \frac{27\,499{,}75 \cdot 3}{45\,882} =$$

$$(=)\ 8 - 0{,}6 \cdot 3\ (=)\ 6{,}2 \text{ ha.}$$

Panašiai ir

$$Q_2 = 20 - \frac{224\,508 - 215\,535{,}75}{33\,236}\,(20 - 15) =$$

$$= 20 - \frac{24\,323{,}75 \cdot 5}{33\,236} =$$

(=) 20 — 1,35 (=) 18,65 ha.

Kvartilinių dydžių suradimas neklasifikuotoj st. eilutėj yra žymiai paprastesnis. Pavyzdžiui galima imti duomenis (lentelė 17a, psl. 180) apie vidutinius derlius iš 1 ha Lietuvos apskrityse. Iš skilties 1 matyti, kad kviečių derliai 23-jose apskrityse svyravo tarp 10,5 ir 18,2 q (21—36,4 centr.) iš 1 ha. Kadangi atsitikimų (apskričių) būta 23 (N = 23), tai kvartiliniai dydžiai Q, kaip ir medianas M gaunama iš karto pagal atitinkamas formulas.

Tuo būdu randame, kad

$$Q_{n,1} = \frac{N + 1}{4} = \frac{23 + 1}{4} = 6, \text{ ir}$$

$$Q_{n,2} = \frac{(N + 1)\,3}{4} = \frac{24 \cdot 3}{4} = \frac{72}{4} = 18, \text{ o}$$

$$M_n = \frac{N + 1}{2} = \frac{24}{2} = 12.$$

Iš lentelės gi matome, kad $x_6 = 11{,}5$ (Ukmergės apskr. derliaus dydis), $x_{12} = 13{,}0$ (Marijampolės apskr. derlius) ir $x_{18} = 13{,}6$ (Biržų apskr. derliaus dydis). Taigi:

1. Jeigu $Q_{n,1} = 6$, o $x_6 = 11{,}5$, tai ir $Q_1 = 11{,}5$ q.
2. Jeigu $Q_{n,2} = 18$, o $x_{18} = 13{,}6$, tai ir $Q_2 = 13{,}6$ q.
3. Jeigu $M_n = 12$, o $x_{12} = 13{,}0$, tai ir $M = 13{,}0$ q.

Kvartiliniai dydžiai Q, visų pirma, parodo ribas, kuriose telpa 50% visos statistinės masės, arba ją atstovaujančios statistinės eilutės dydžių. Šiuo atveju, pav., kai $Q_1 = 6{,}20$, o $Q_2 = 18{,}65$, mes jau žinome, kad iš visų žemės ūkių Lietuvoje 50% sudaro ūkiai nuo 6,20 ligi 18,35 ha. Be to, žinodami, kad mažiausias ūkis turi 1 ha, o patsai vidurinis (medianas) 10,86 ha, taip pat jau žinome, kad ¼ ūkių (25%) yra nuo 1 ligi 6,60 ha ir antra tiek (taip pat 25%) nuo 6,20 ligi 10,86 ha. Taip pat jau

žinome, kad 25% ūkių turi 10,86—18,65 ha ir 25% ūkio yra stambesni negu 18,65 ha. Taigi jau 75% visos masės iš kvartilinių dydžių gana griežtai apibrėžiami ir tik 25% lieka nepakankamai ištirti.

Kvartiliams (kaip ir medianui) patiektoji formula atitinka F. H. Harper'io apibrėžtas sąlygas (žiūr. jo „Elements of Practical Statistics", p. 146 ir kt.). Daugelyj statistikos vadovėlių, tačiau, medianui ir kvartiliams surasti apibendrintų formulų visai neduodama, o kai kurie autoriai kvartilių formulas apibrėžia kitaip. Pav., Sergant Florance (The Statistical Method in Economics and Political Science p. 496) rekomenduoja: $Q_1 = \frac{N}{4} + \frac{1}{2}$ ir $Q_2 = \frac{3N}{4} + \frac{1}{2}$. Kai kurie kiti autoriai siūlo pirmąjį kvartilį (Q_1) tiesiog iš $\frac{N}{4}$, o antrąjį (Q_2) iš $\frac{3N}{4}$, be jokių priedėlių. 1937/8 mt. statistikos seminaro dalyvis, matematikas Jurskis, savo referate rekomendavo kvartiliams formulas, atitinkančias griežtesnio g e o m e t r i n i o matavimo sąlygas $\left(Q_1 = \frac{N+3}{4} \text{ ir } Q_2 = \frac{3N+1}{4}\right)$.

Statistikos praktikai, tačiau, F. H. Harper'io apibendrintoji kvartilių formula yra tikslingiausia. Kvartiliai statistikoj savo esmėje yra vad. Galton'inės klasifikacijos varijantas. Kvartiliais, kai jie imami kartu su medianu ir pačiais kraštutiniais dydžiais, visa statistinė masė suskaidoma į keturias lygias grupes, ir kvartiliniai dydžiai terodo tik kiekvienos tokios grupės intervalo aukštutinę ar žemutinę ribą. Tais atvejais, kai hipotetinis (pagal formulą išskaičiuotas) kvartilinis dydis (ar medianas) sutampa su empirinio stebėjimo keliu gautuoju dydžiu, — gretimųjų grupių intervalai atsiremia į empirinį dydį; o jeigu nesutampa, — jie parodo tik g a l i m o arba hipotetinio dviejų gretimų dydžių atsirėmimo (arba susiekimo) ribą. Iš kitos pusės, statistinio stebėjimo keliu gautieji empiriniai dydžiai, bet kuriuo statistikos metodu klasifikuojami, visuomet palieka n e s k a i d o m i skaitomieji v i e n e t a i. Taigi, statistikos praktikoj negali būti atsitikimo, kad išeitų kokia nors kolizija tarp hipotetinių kvartilių ar hipotetinio mediano iš vienos pusės ir atitinkamų skaitomųjų vienetų statistinėj eilutėje: p a g a l b e t k u r i ą f o r m u l ą g a u t i e j i k v a r t i l i n i a i d y d ž i a i y r a t i k r o d i k l i a i, k u r i a i s t e n k a v a d o v a u t i s b e i e š k a n t s t a t i s t i n ė j e m a s ė j e e m p i r i n i ų d y d ž i ų. Jeigu formulos patiektas kvartilio rodiklinis dydis atitinka empiriško stebėjimo keliu rastąjį skaitomojo vieneto dydį x_i su absoliutiniu griežtumu — gerai; bet jeigu tokio dydžio statistinėje masėje nėra, tenka imti patį artimiausiąjį dydį. Jei gi atsitiktų, kad kurio nors kvartilio (arba mediano) rodiklinį dydį atitinkančių empirinių dydžių statistikos eilutėje yra ne vienas, o keletas ar keliasdešimtis, tai vienodo dydžio skaitomųjų vienetų grupė tektų perskirti ir prie žemesnėsės (arba aukštesnėsės) kvartilinės klasės priskirti tik tiek, kiek reikalinga, kad sudaryti statistinį dalinį iš $\frac{1}{4}$ skaitomųjų vienetų. Likusieji gi tokio pat didumo skaitomieji vienetai turėtų būti priskirti prie sekančios kvartilinės grupės.

Šitose sąlygose apibendrintoji kvartilių (ir mediano) formula įgauna statistinę prasmę tik tiems atvejams, kai skaitomieji vienetai yra išri-

kiuoti sulig vyresniškumu ne paskirai, o grupėmis su mažesniais ar didesniais intervalais. Šiose sąlygose pagal formulą nustatytas kvartilinio (ar medianinio) dydžio eilinis n ir jį atitinkantis dydis x_i yra tik orientacinis skaičius, kuriuo vadovaujantis galima vienu ar kitu būdu surasti ir tikrai esantį, ar bent jam artimesnį, empirinį dydį. Jeigu empiriniai dydžiai yra jau sugrupuoti į dydžio grupes su apibrėžtais intervalais, pagal formulą gautasis kvartilio (ar mediano) rodiklinis dydis palieka galutinai, o reikalingoji atitinkamos dydžio grupės empirinių vienetų dalis priskiriama prie vienos ar kitos kvartilinės grupės proporcionaliai pagal gautąjį kvartilio (ar mediano) hipotetinį (rodiklinį) dydį tiesinės interpoliacijos keliu.

Savaime suprantama, kad tokiose aplinkybėse interpoliacijos keliu gautasis dydis gali nesutapti su atitinkančio eilinio nr. faktinuoju arba empirinių dydžių x_i. Tačiau paklaida negali būti didesnė kaip skaitomajam vienetui apibrėžto dydžio mato vienetas. Jeigu, pav., ūkius skaidyti į dydžio grupes po 1 ha, tai gali atsitikti, kad viena kvartilinė grupė pasibaigs ūkiu, kurio tikrasis dydis sieks 3,99 ha, o sekanti kvartilinė grupė, kuri teoriškai apims sekantį ūkį, bent kiek didesnį kaip 3,99 ha, praktiškai galės gauti sekantį ūkį su 4,00 arba paskutinį su 4,99 ha (jei ūkių tarp 4,01—4,98 ha visai nebūtų). Ši maksimalinė paklaida, tačiau, tikrovėje sumažėja ligi ½ sveikojo mato vieneto, jei išeiti iš bet kurios formulos duodamo hipotetinio kvartilių arba mediano dydžio. Nes ar mes pirmojo kvartilio ieškosime pagal rodiklinį dydį, gautą iš N : 4, ar prie gautojo dalmens pridėsime 0,5 — galutinieji rezultatai negalės pasikeisti daugiau kaip 1—0,5, — jeigu kvartilinis dydis bus interpoliuotas.

Tuo būdu apibendrintos kvartilinių dydžių formulos uždavinys yra technikinis, ir net ideališkiausia formula nuo paklaidos apsaugoti negali. Dėl to statistikos praktikai tenka pripažinti tinkamiausia tokia formula, kuri, vidutinės paklaidos nepadidindama, lengviausia pritaikoma.

Tuo būdu kvartiliai įgalina visą statistinę masę suskirstyti jos vienetų didumu į keturias dydžio klases ir kiekvienos klasės intervalus apibrėžti. Bet iš kvartilių galima apytikriai spręsti ir apie statistinės masės pasiskaidymo pobūdį, būtent: ar st. masės medianui artimųjų dydžių, sudarančių visos masės 50%, pasiskaidymas yra simetriškas, ar nesimetriškas. O tai yra svarbu žinoti ypač dėl to, kad iš kvartilinių dydžių santykio su medianu galima jau iš dalies spręsti ir apie dispersijos pobūdį, t. y. ar st. masės pasiskaidymas artimesnis vad. normalios dispersijos tipui, ar ne. Visa tai parodo atitinkami statistiniai skaičiai: a) asimetrijos koeficientas ir b) dispersijos koeficientas.

Metodų asimetrijos ir dispersijos koeficientams gauti yra nemaža. Šioj, tačiau, vietoj pravartu susipažinti su kvartilinių dydžių ypatybėmis, taigi ir su kvartiliniais asimetrijos ir dispersijos koeficientais.

Asimetrijos koeficientas aplamai turi parodyti, kurios rūšies — smulkesniųjų ar stambesniųjų — st. ma-

sės vienetų yra daugiau ir koks yra jų lyginamasis svoris. Jei sutarsime kvartilinį asimetrijos koeficientą žymėti a_q, tai surasti jį galėsime arba betarpiškai iš kvartilių pagal formulą:

$$a_q = \frac{Q_2 + Q_1 - 2M}{Q_2 - Q_1} \text{ arba} \qquad \text{(XXIa)}.$$

$$a_q = \frac{q_2 - q_1}{q_2 + q_1} = \frac{(Q_2 - M) - (M - Q_1)}{(Q_2 - M) + (M - Q_1)} \qquad \text{(XXIb)}.$$

Lengva įsitikinti, kad abidvi formulos yra tapatingos, t. y., kad

$$a_q = \frac{Q_2 + Q_1 - 2M}{Q_2 - Q_1} = \frac{q_2 - q_1}{q_2 + q_1}.$$

Pavyzdiniam skaičiavimui imkime tuos pačius 1930 mt. žemės ūkio surašymo duomenis. Jau anksčiau esame nustatę, kad:

$$M = 10{,}86,\ Q_1 = 6{,}20 \text{ ir } Q_2 = 18{,}65 \text{ ha}.$$

Taigi kvartilinis asimetrijos koeficientas pagal formulą:

$$a_q = \frac{Q_2 + Q_1 - 2M}{Q_2 - Q_1} = \frac{18{,}65 + 6{,}20 - 2 \,.\, 10{,}86}{18{,}65 - 6{,}20} =$$

$$= \frac{24{,}85 - 21{,}72}{12{,}45} = \frac{3{,}13}{12{,}45} = 0{,}25.$$

Visai tą patį $a_q = 0{,}25$ duoda ir formula XXIb, nes: kai

$$M = 10{,}86,\ Q_1 = 6{,}20 \text{ ir } Q_2 = 18{,}65 \text{ ha},$$

tai

$$q_1 = M - Q_1 = 10{,}86 - 6{,}20 = 4{,}66 \text{ ha},$$
$$q_2 = Q_2 - M = 18{,}65 - 10{,}86 = 7{,}79 \text{ ha}.$$

Vadinasi ir asimetrijos koeficientas pagal formulą

$$a_q = \frac{q_2 - q_1}{q_2 + q_1} = \frac{7{,}79 - 4{,}66}{7{,}79 + 4{,}66} = \frac{3{,}13}{12{,}45} = 0{,}25.$$

Šiuo atveju gautas teigiamas asimetrijos koeficientas rodo, kad aritmetinis vidurkis turėtų būti didesnis negu medianas, t. y. didesnis kaip 10,86 ha. Iš ankstyvesniųjų duomenų (psl.

170) mes žinome, kad a. vidurkis siekia net 15.03 ha. O jeigu taip, tai dominantas turėtų būti mažesnis už 10,86 ha.

Tuo būdu praktiška kvartilinio asimetrijos koeficiento nauda yra ta, kad st. eilutei apibūdinti galime prasmingai panaudoti vad. matematinį metodą vidurkiams gauti, — jeigu svertiniam vidurkiui gauti sunkenybių būtų.

Dispersijos koeficientas privalo atvaizduoti st. eilutės vienetų skirtumus, lyginant su vidurkiu. Kvartilinis dispersijos koeficientas, žinoma, tegali atvaizduoti tokį santykį su medianu. Gaunamas jis iš kvartilių vidurkio pagal formulą:

$$d_q = \frac{\frac{1}{2}(Q_2 - Q_1)}{M} \qquad \text{(XXIIa).}$$

$$d_q = \frac{\frac{1}{2}(q_2 + q_1)}{M} \qquad \text{(XXIIb).}$$

Mūsų pavyzdyje gausime:

$$d_q = \frac{\frac{1}{2}(Q_2 - Q_1)}{M} = \frac{\frac{1}{2}(18{,}65 - 6{,}20)}{10{,}86} =$$

$$= \frac{0{,}5 \,.\, 12{,}45}{10{,}86} (=) \frac{6{,}23}{10{,}86} (=) \; 0{,}58.$$

Patikrinę pagal formulą XXIIb gauname:

$$d_q = \frac{\frac{1}{2}(q_2 + q_1)}{M} = \frac{0{,}5\,(7{,}79 + 4{,}66)}{10{,}86} =$$

$$= \frac{0{,}5 \,.\, 12{,}45}{10{,}86} (=) \frac{6{,}23}{10{,}86} (=) \; 0{,}58.$$

Tuo būdu randame, kad kvartilinis dispersijos koeficientas sudaro net 58% medianinio dydžio. Vadinasi, dar neatsižvelgę kraštutinių nukrypimų, o tik paėmę 50% vienetų iš vidurio masės, sutinkame jau gana didelių skirtumų.

Tokiose aplinkybėse net negalima galutinai spręsti, ar tyrinėjamos masės dispersija dar galima laikyti bent artutine normaliai dispersijai.

Aplamai, kvartiliniais asimetrijos ir dispersijos koeficientais, arba kvartiliniais dydžiais, galima pasitenkinti tik tais atvejais, kai šie koeficientai, apskritai imant, išeina nedideli, vadinasi, kai yra ženklas, kad bent 50% visos statistinės masės nedaug kuo skiriasi nuo medianinio dydžio. Tokiu atveju likusieji 50% visos masės kad ir skirtųsi žymiau nuo mediano, bet jau negalėtų nulemti dispersijos pobūdžio (tą pobūdį vaizduojančių asimetrijos ir dispersijos koeficientų). Tokį asimetrijos atsitikimą mes galime rasti pavyzdyje iš lentelės psl. 180 apie derlius iš 1 ha. Šiuo atveju, kai M = 13,0, Q_1 = 11,5 ir Q_2 = 13,6, kvartilinis dispersijos koeficientas pagal formulą išeis 0,08 arba 8%, nes:

$$d_q = \frac{\frac{1}{2}(13{,}6 - 11{,}5)}{13{,}0} = \frac{1{,}05}{13{,}00} \; (=) \; 0{,}08 \text{ arba } 8\%$$

Šiose aplinkybėse likusieji 50% dispersijos esmės iškreipti negalėtų.

c) Paprastieji arba aritmetiniai nukrypimai bei asimetrijos ir dispersijos koeficientai.

Griežtesnis statistinių masių arba statistikos eilučių analizis yra reikalingas asimetrijos ir dispersijos koeficientų, kurie rodytų visos st. masės vienetų susikaupimo bei skirtumų bent apytikrį vaizdą. Čia ir vėl kai kada galima pasikliauti mažiau griežtais, arba tik artutiniais, skaičiavimais, bet gali būti reikalingi ir griežtesni st. analizio metodai. Ypač visais tais atvejais, kai turima duomenis iš reprezentacinio stebėjimo, arba kai ištisinio stebėjimo duomenys norima smulkiau išnagrinėti reprezentaciniu metodu, — negalima pasikliauti nei kvartiliniais dydžiais, nei bet kuriais vadinamais vidutiniais arba paprastais aritmetiniais koeficientais. Tokiais atvejais yra reikalingi griežtesni matematinio analizio metodai.

Kai duomenys yra gauti ištisiniu stebėjimu, arba iš labai didelio reprezentantų skaičiaus, ir kai, be to, tyrinėjamos masės vienetų skirtumai nėra dideli, ligi tam tikro laipsnio galima pasikliauti ir paprastaisiais skirtumais nuo aritmetinio vidurkio vad. v i d u t i n i a m dispersijos koeficientui gauti. Tam reikalui imami visų statistinės eilutės dydžių skirtumai nuo arit-

metinio vidurkio, sumuojami neatsižvelgiant į tai, ar tie skirtumai yra teigiami ar neigiami, ir gautoji suma daloma statistinės eilutės narių skaičiumi. Tuo būdu gautasis skirtumų vidurkis atstovauja jau visą statistinę eilutę, o jo santykis arba lyginys su vidurkiu parodo vadinamą vidutinę, arba aritmetinę, dispersiją. Šitaip išskaičiuoto vidutinio nukrypimo ribose apie vidurkį sutelpa 50% visų statistinės eilutės dydžių. Suradus tai pačiai statistikos eilutei medianą ir turint aritmetinį vidurkį ir vidutinį nuo jo nukrypimą galima išskaičiuoti ir asimetrijos koeficientą.

Pirmiausia tenka išskaičiuoti vad. vidutinį nukrypimą. Jei sutarsime tokį nukrypimą žymėti raide v, tai jam gauti iš x-inės eilutės skirtumų nuo vidurkio $\bar{x}$ turėsime formulą:

$$v = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \ldots + |x_i - \bar{x}| + \ldots + |x_n - \bar{x}|}{n} = \frac{\Sigma |x_i - \bar{x}|}{n} \qquad \text{(XXIIIa)}$$

Jeigu x-inės eilutės dydžiai pasikartoja nevienodai, o kiekvienas turi savojo gausumo skaičius y_i, tai ta pati formula atrodys šitaip:

$$v = \frac{|x_1 - \bar{x}|y_1 + |x_2 - \bar{x}|y_2 + \ldots + |x_i - \bar{x}|y_i + \ldots + |x_n - \bar{x}|y_n}{y_1 + y_2 + \ldots + y_i + \ldots + y_n} = \frac{\Sigma |x_i - \bar{x}|\, y_i}{\Sigma y_i} \qquad \text{(XXIIIb).}$$

Tenka įsidėmėti, kad XXIII formuloj $x_i - \bar{x}$ reiškiniai nėra įsikliausti $(x_i - \bar{x})$, o tik statmeniniais brūkšniais atžymėti: $|x_i - \bar{x}|$. Brūkšniai parodo, kad $x_i - \bar{x}$ skirtumai visais atvejais rašomi teigiamais skaičiais[6]. O tai reiškia, kad jeigu $x_i < \bar{x}$, tai skirtumas, gaunamas pagal formulą

$$d_i = \bar{x} - x_i. \quad \text{O } d_i = x_i - \bar{x}$$

betarpiškai gaunamas tik tuo atveju, jei $x_i > \bar{x}$.

---

[6] Pav., $|-3| = |3| = 3$. Ženklas $|a|$ rodo, kad imamas tikrojo skaičiaus *a* absoliūtus didumas (modulis), kuris visuomet nustato teigiamą skaičių.

Be to, $(x_i - \bar{x})$ skirtumams žymėti sutarę raidę *d*, ir XXIII formulos variantus galima rašyti:

$$v_a = \frac{\Sigma d_i}{n} \quad \text{ir} \quad v_b = \frac{\Sigma d_i y_i}{\Sigma y_i}.$$

Savaime suprantama, kad rezultatai bus gauti tie patys, jei vidurkis ir skirtumai bus skaičiuojami pasikliaujant konstanta *a*. Šiuo atveju vid. nukrypimui *v* skaičiuoti formula suderinama su aritmetinių vidurkių skaičiavimo formula VIII.

Turint vidutinį a. nukrypimą, aritmetinis dispersijos koeficientas d_a gaunamas pagal formulą:

$$d_a = \frac{v}{m} = \frac{v}{\bar{x}} \quad \text{arba} \quad d_a\% = \frac{v.\,100}{m} = \frac{v.\,100}{\bar{x}} \qquad \text{(XXIV)}.$$

Aritmetinis gi asimetrijos koeficientas a_a gaunamas pagal formulą:

$$a_a = \frac{m - D}{v} \,(=)\, \frac{3\,(m - M)}{v} =$$

$$= \frac{\bar{x} - D}{v} \,(=)\, \frac{3\,(\bar{x} - M)}{v} =$$

$$= \frac{\bar{x} - D}{\frac{1}{n}\,\Sigma\,|\,x_i - \bar{x}\,|} = \frac{3\,(\bar{x} - M)}{\frac{1}{n}\,\Sigma\,|\,x - \bar{x}\,|} \qquad \text{(XXV)}.$$

Skaičiavimo pavyzdžiu galime paimti kviečių derliaus duomenis iš 1932—1934 mt. laikotarpio. Išrikiavę tuos duomenis pagal vidutinį derlių kvintalais iš 1 ha gauname šitokią lentelę (žiūr. p. 207):

Kai aritmetinis vidurkis $\bar{x} = 12{,}96$ q, bendroji $|\,x - \bar{x}\,|$ skirtumų d_i suma $\Sigma d_i = 30{,}24$ ir x_i eilutės narių skaičius $n = 23$, tai vidutinis nukrypimas *v* pagal XXIII formulą:

$$v = \frac{\Sigma d_i}{n} = \frac{30{,}24}{23} \,(=)\, 1{,}32 \text{ q}.$$

Tuo būdu aritmetinis dispersijos koeficientas gaunamas

$$d_a = \frac{v}{\bar{x}} = \frac{1{,}32}{12{,}96} \,(=)\, 0{,}1 \text{ arba } 10\%.$$

Lentelė 17.

KVIEČIŲ DERLIAI IŠ 1 HA Q 1932/34 MT.

n	Apskritys	Derliai iš 1 ha	Skirtumai nuo vidurkio $\bar{x} = 12,96$
		x_i	$\lvert x_i - \bar{x} \rvert = d_i$
1	Telšiai	10,50	2,46
2	Utena	11,00	1,96
3	Seinai	11,20	1,66
4	Trakai	11,30	1,66
5	Zarasai	11,40	1,56
6	Ukmergė	11,50	1,46
7	Kaunas	11,60	1,36
8	Alytus	11,90	1,06
9	Kretinga	11,90	1,06
10	Šakiai	12,60	0,36
11	Panevėžys	12,70	0,26
12	Marijampolė	13,00	0,04
13	Tauragė	13,30	0,34
14	Rokiškis	13,30	0,34
15	Kėdainiai	13,30	0,34
16	Raseiniai	13,40	0,44
17	Mažeikiai	13,60	0,64
18	Biržai	13,60	0,64
19	Šiauliai	13,70	0,74
20	Klaipėda	14,70	1,74
21	Vilkaviškis	15,10	2,14
22	Šilutė	15,70	2,74
23	Pagėgiai	18,20	5,24
		$29,85 = \Sigma x_i$	$\Sigma d_i = \frac{30,24}{23} = 1,315$

Yra taisyklė, kad vidutinio nukrypimo ribose apie vidurkį turi tilpti 50% visų st. eilutės dydžių. Galime tuojau patikrinti. Kai vidurkis $\bar{x} = 12,96$, o vidutinis nukrypimas $v = 1,32$, tai intervalas k_i

$$k_v = \bar{x} \pm v = 12,96 \pm 1,32 =$$
$$= 11,64 \longrightarrow 14,28.$$

Iš lentelės galime įsitikinti, kad šiose ribose sueina visi x_i pradedant x_8 (Alytaus apskr.) ir baigiant x_{19} (Šiaulių apskr.), taigi iš viso 12 iš 23.

Toj pačioj lentelėj lengvai surandamas ir medianas. Kadangi n = 23, tai formula:

$$M_n = \frac{n+1}{2} = \frac{23+1}{2} = 12.$$

Dvyliktąjį gi x_i atstovauja Marijampolės apskr. derlius — 13,00 q. Taigi M = 13,00. O turėdami M, $\bar{x}$ ir v, surandamas ir aritmetinis asimetrijos koeficientas:

$$a_a = \frac{3(\bar{x} - M)}{v} = \frac{3(12{,}96 - 13{,}00)}{1{,}32} = \frac{-0{,}04 \cdot 3}{1{,}32} =$$

$$= \frac{-0{,}12}{1{,}32} (=) -0{,}08.$$

d) Normalioji dispersija ir vidurinis kvadratingasis nukrypimas.

Vad. normalios dispersijos ideališką tipą atstovauja žinomoji binomo koeficientų dispersija. Socialekonominio gyvenimo reiškiniuose normaliai artutinė dispersija sutinkama kiekvieną kartą, kai mažiau ar daugiau lygiareikšmiai pažymiai pakankamai didelėj statistinėj masėj turi vienodus šansus vienaip ar kitaip pasireikšti.

Pirmiausia prisiminkime iš algebros charakteringąjį binomo koeficientų pasiskaidymą. Žinomoji binomo formula

$$(a+b)^n = a^n + \binom{n}{1} a^{n-1}b + \binom{n}{2} a^{n-2}b^2 + \ldots + b^n$$

aiškiai parodo, kad dviejų galimų įvykių, arba dviejų pažymių, kombinacijų skaičius auga geometrinėj progresijoj. Tuo tarpu kiekvienos iš viso galimos tų dviejų pažymių kombinacijos tikimybė turi pasireikšti didesne ar mažesne trupmena, jei operuojama natūriniais skaičiais. Bet kuris binomo koeficientas $\binom{n}{k}$, kuris žymimas dar ir $(n)_k$ arba nC_k, gaunamas pagal formulą:

$$\binom{n}{k} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{1 \cdot 2 \ldots (k-1)\,k} = \frac{n!}{k!\,(n-k)!}$$

$$= (n)_k = nC_k \qquad \text{(XXVI).}$$

Čia $k! = 1 \cdot 2 \cdot 3 \ldots k$.

Pav., jei imsime monetą ir mesime tik vieną kartą, tai gali atsiversti vytis (A) arba skaičius (B). Vadinasi, šiuo atveju $\binom{n}{k} = 1$, nes iš dviejų iš viso galimų atvejų praktiškai tėra galimas tik vienas (A a r b a B). Jeigu gi $n = 1$, vadinasi, ir $A + B = 1$, tai bet kurio (A ir B) nario tikimybės koeficientas $\binom{n}{k} : 2^1 = \frac{1}{2}$. Tačiau, jeigu monetos metimą pakartosime, tai galimų kombinacijų jau bus 4: nes du kartu metant pinigą gali atsiversti arba $AB = \frac{1}{4}$, arba $BA = \frac{1}{4}$ arba $AA = \frac{1}{4}$ arba $BB = \frac{1}{4}$.

$$\frac{1}{4} + \frac{1}{4} + \left(\frac{1}{4} + \frac{1}{4}\right) = \frac{1}{4} + \frac{1}{4} + \frac{2}{4} = 1.$$

Tris kartus mesdami monetą galime gauti šitokias kombinacijas: AAA, AAB, ABA, BAA, BBB, BBA, BAB ir BBB. Žodžiu, gausime jau aštuonias kombinacijas, iš kurių dvi (3A ir 3B) tegalės pasireikšti tik viename variante (vadinasi, jų tikimybės koeficientas $\binom{n}{k} : 2^2 = \frac{1}{8}$, ir dvi kombinacijos (2A + 1B ir 2B + 1A) turės po tris variantus (kiekvienos iš jų tikimybė $\binom{n}{k} : 2^3 = \frac{3}{8}$).

Keturis kartus metant monetą bus jau 16 galimų kombinacijų ($2^4 = 16$):

4A	$^1/_{16}$
3A + 1B	$^4/_{16}$
2A + 2B	$^6/_{16}$
1A + 3B	$^4/_{16}$
4B	$^1/_{16}$
16k	$\frac{16}{16} = 1$

Aplamai, jeigu $a = b$ ir $a = 1$, tai binomas $(a + b)^n$ virsta $(1 + 1)^n = 2^n$. Binomo koeficientai bet kuriam n laipsniui įgauna reikšmes, kurias galime sugrupuoti į vadinamąjį aritmetiškąjį trikampį:

n = 0					1						
1				1		1					
2			1		2		1				
3		1		3		3		1			
4	1		4		6		4		1		
5		1		5		10		10		5	1
6	1		6		15		20		15	6	1
.					.....						
.					.....						
.					.....						
10	1	10	45	120	210	252	210	120	45	10	1
.					.....						
.					.....						

Ši binomo koeficientų lentelė duodasi tęsiama ir tęsiama be galo, ir kiekvienas aukštesnio laipsnio eilutės narys gaunamas iš žemesniojo laipsnio dviejų gretimųjų narių sumos.

Statistikui binomo koeficientų pasiskaidymas turi reikšmę, visų pirma, dėl charakteringų jo koeficientų dispersijos pobūdžio. Dalykas toks, kad binomo koeficientų pasiskaidymas bet kurio laipsnio eilutei visais atvejais yra griežtai simetriškas, be to, juo aukštesnio laipsnio eilutė — vadinasi, juo daugiau įvairių galimų kombinacijų, tuo smulkesnės gaunamos kiekvienos atskirai paimtos kombinacijos tikimybių trupmenos. Pav., antrojo laipsnio binomo koeficientų eilutėj mažiausia tikimybių trupmena yra $\frac{1}{4}$, o didžiausia $\frac{1}{2}$. Tuo tarpu ketvirtojo laipsnio binomo koeficientus atitinkančių tikimybių eilutėj mažiausia trupmena $\frac{1}{16}$, o didžiausia — $\frac{4}{16}$. Dešimtojo gi laipsnio (2^{10}) eilutėj didžiausioji trupmena sumažėja ligi $\frac{252}{1024} = 0{,}246$, t. y. išeina mažesnė už antrojo laipsnio mažiausiąją trupmeną, kai mažiausiųjų tikimybinių koeficientų net bendroji suma tesudaro tik $\frac{1}{512}$, o pats vienas mažiausias koeficientas yra $\frac{1}{1024}$.

Be to, binomo koeficientų pasiskaidymas dar turi ir antrą, statistikos eilučių dispersijos atžvilgiu labai reikšmingą ypa-

tybę: bet kurio laipsnio binomo koeficientų eilutėj gausingiausia yra atstovaujamas pats vidurinis dydis, t. y. medianas. O iš vidurkių teorijos mes jau žinome, kad jeigu dominantas (D) atitinka medianą (M), tai tuo pačiu nulemiamas ir aritmetinis vidurkis ($\bar{x}$). Vadinasi, gaunama skaičių eilutė, kur $\bar{x} = M = D$, t. y. visi trys skaičių eilutės charakteringieji dydžiai yra tolygūs.

Šios rūšies skaičių arba st. eilutės dydžių pasiskaidymas priklauso vadinamojo n o r m a l a u s pasiskaidymo, arba n o r m a l i o s d i s p e r s i j o s. Palieka tiktai apibrėžti šitos normalios dispersijos g a l i m ų, nors ir nebūtinų, nukrypimų ribas, o taip pat dėsnius, kurių priklauso apibrėžtose ribose normalios dispersijos galimi nukrypimai bei varijantai. Šiuos apibrėžimus turėdami, turėsime ir st. eilučių, taigi statistinių masių, matavimo normas.

Jau anksčiau buvo minėta, kad st. eilučių dispersijos pobūdžiui apibrėžti pirmoj eilėj vartojama ir vartotina dvejetas statistinių mastelių: dispersijos koeficientas ir asimetrijos koeficientas. Pati vad. normali dispersija ir jos normų dėsniai bei patys koeficientai yra matuojami pasikliaujant vad. v i d u r i n i u k v a d r a t i n g u o j u n u k r y p i m u, kurio paskirtis, aplamai imant, yra ta pati kaip ir vad. vidutiniojo aritmetinio nukrypimo v (žiūr. psl. 205). Skirtumas, tačiau, tas, kad vad. vidurinis kvadratingasis nukrypimas, paprastai žymimas mažąja sigma (σ), yra betarpiškai susietas su tikimybėmis ir duoda matematiškai pagrįstą mastą st. eilučių dispersijos normoms apibrėžti ir nustatyti, ar duotosios st. eilutės dispersija yra normali ar nenormali.

V i d u r i n i s k v a d r a t i n g a s i s n u k r y p i m a s σ skaičiuojamas sulig skirtumais nuo vidurkio pagal formulą

$$\sigma = \sqrt{\frac{\Sigma \triangle_i^2}{N}} = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{N}} \qquad \text{(XXVIIa)},$$

jeigu st. eilutė yra paprasta (st. eilutės x dydžiai nepasikartoja), arba pagal formulą

$$\sigma = \sqrt{\frac{\Sigma \triangle_i^2 y_i}{\Sigma y_i}} = \sqrt{\frac{\Sigma (x_i - \bar{x})^2 y_i}{\Sigma y_i}} \qquad \text{(XXVIIb)}.$$

jei x_i dydžiai turi skirtingus svorio koeficientus (y_i).

Delta (△) šioj formuloj atstovauja skirtumus nuo aritmetinio vidurkio. Tuo būdu st. eilutėj bet kuris $x_i - \bar{x} = \triangle_i$. Vadinasi, patsai kvadratingas nukrypimas kvadrate (σ^2) gaunamas pagal formulą:

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n} =$$

$$= \frac{1}{n} \Sigma (x_i - \bar{x})^2.$$

Sutarus st. eilutės dydžių skirtumus nuo vidurkio žymėti △ ir įvedus lygybę: $\triangle_i = x_i - \bar{x}$, ta pati formula atrodys paprasčiau:

$$\sigma^2 = \frac{\triangle_1^2 + \triangle_2^2 + \triangle_3^2 + \ldots + \triangle_n^2}{n} = \frac{\Sigma \triangle_i^2}{n}, \text{ arba}$$

$$\sigma^2 = \frac{\triangle_1^2 y_1 + \triangle_2^2 y_2 + \triangle_3^2 y_3 + \ldots + \triangle_n^2 y_n}{y_1 + y_2 + y_3 + \ldots + y_n} = \frac{\Sigma \triangle_i^2 y_i}{\Sigma y_i}.$$

Svarbu įsidėmėti, kad vidurinis kvadratingasis nukrypimas σ skaičiuojamas iš skirtumų nuo aritmetinio vidurkio. O žinome, kad skirtumai nuo aritmetinio vidurkio yra visais atvejais m a ž i a u s i, taigi ir šių skirtumų kvadratų suma yra mažiausia (mažesnė už bet kurių kitų skirtumų kvadratų sumą). Jau ši viena aplinkybė vidurinį kvadratingąjį nukrypimą (σ) daro visiems galimiems atvejams vienodai patikimu rodikliu. Be to, vad. n o r m a l i o s d i s p e r s i j o s st. eilutės vid. k v a d r a t i n g a s i s n u k r y p i m a s lyginant jį su vidurkiu, bus tuo mažesnis, kuo tiriamoji masė didesnė. Pagaliau, griežtai arba ideališkai normalioj dispersijoj visi st. eilutės dydžiai (x-inės eilutės variantai) apie tos pačios eilutės vidurkį pasiskaido tam tikra, visiems atvejams vienoda, tvarka, būtent:

$\bar{x} \pm 1\sigma$	ribose	sueina	apie	68,27%	visų	dydžių
$\bar{x} \pm 2\sigma$	„	„	„	95,45%	„	„
$\bar{x} \pm 3\sigma$	„	„	„	99,73%	„	„

Tai yra vadinamoji t r i j ų s i g m ų t a i s y k l ė, kuri, mokamai ir atsargiai naudojama, statistikos praktikoj gali daug pa-

tarnauti. Juo labiau, kad ši trijų sigmų taisyklė, kaip yra įrodę ypač kai kurie rusų ir kt. matematikai ir statistikai (pirmoj eilėj Čebyšev, Markov, o taip pat O. Anderson, A. Fisher ir kt.), su tam tikromis pataisomis galioja bet kurio tipo dispersijai, vadinasi, ir asimetriškai st. eilutei. Pagal Čebyševo teoremą, tikimybė, kad bet kurios rūšies ar tipo st. eilutėj bet kuris dydis randasi tam tikrose ribose, duodasi apibrėžiama šitaip:

	nenormalios dispersijos st. eilutėj	normalios disp. st. eilutėj
$\bar{x} \pm 1\sigma$	ribose x_i sutelpa daugiau kaip 0	apie 68,27%
$\bar{x} \pm 2\sigma$	„ „ „ „ „ 75%	„ 95,45%
$\bar{x} \pm 3\sigma$	„ „ „ „ „ 88,9%	„ 99,73%

Vadinasi, net ir nenormalios dispersijos bei asimetriškoj st. eilutėj aritmetinis vidurkis su $\pm 3\sigma$ pataisa ($\bar{x} \pm 3\sigma$ intervalo ribose) atstovauja nemažiau kaip 89% visos st. masės, normalios gi dispersijos eilutėj net daugiau kaip 99,7%, t. y. praktiškai beveik visą 100%. O normalios dispersijos ideališką tipą atstovauja binomo koeficientų eilutės vad. aritmetinis trikampis, kuris, koordinatų sistemoj grafiškai atvaizduotas, duoda charakteringą bei taisyklingą kreivę su griežtai vienodais kilimo ir kritimo laipsniais abiejose koordinatų ašies pusėse. Bet ir paprastoj st. lentelėj binomo koeficientus padarius st. eilutės skirtingųjų x_i dydžių gausumo koeficientais y_i, o pačius x_i pareiškus galimų potyrių sekos n skaičiais, gausime vad. kvadratingąjį nukrypimą (σ), kurio pagalba vad. trijų sigmų taisyklės esmę ir prasmę galėsime suvokti.

Pavyzdžiui, vėl prisiminkime visas teoretiškai galimas kombinacijas atsiversti A ar B pusėms (skaičiui ar vyčiai) metant, tarkime, 10 kartų. Kadangi šiuo atveju $A = B$ ir $A = 1$, tai $(A + B)^n = (1 + 1)^{10} = 2^{10} = 1024$. Bet kurio n atvejo 1024 galimybių sumoj pasireikš svorio koeficientu pagal jau žinomą binomo koeficientų formulą:

$$\binom{n}{k} = \frac{n!}{(n-k)!\,k!}$$

Statistikos lentelę šiam reikalui galima ir taip sudaryti:

Lentelė 18.

GALIMŲ A IR B ĮVYKIŲ SKAIČIUS LYGIOSE SĄLYGOSE.

Metimų eilinis nr. (n)	Jeigu n = 10 ir A atsivers kartų (k)	tai B teks (n — k) (x_i)	Atsiverst A ir B galimų atvejų skaičius (y_i)
0.	10	0	1
1.	9	1	10
2.	8	2	45
3.	7	3	120
4.	6	4	210
5.	5	5	**252**
6.	4	6	210
7.	3	7	120
8.	2	8	45
9.	1	9	10
10.	0	10	1
		$\Sigma x_i = 55$	$\Sigma y_i = 1024$

Čia turime st. lentelę, iš kurios matome, kad atsiverst A ar B galimų atvejų didžiausią skaičių (252) atstovauja n = 5, t. y. pats vidurinis dydis arba medianas (M), kuris šiuo atveju visai atitinka paprastąjį ir svertinį aritmetinį vidurkį.

Dabar pamėginkime patikrinti ir kartu surasti vidurinį kvadratingąjį nukrypimą. Šiam reikalui x_i eilutėj duosime n skaičių, o x_i dydžių svorio koeficientams y_i tiesiog paimsime binomo koeficientų eilutės dydžius (žiūr. p. 215).

Kadangi šiuo atveju x_i eilutės aritmetinis vidurkis $\bar{x} = 5$, o pagal trijų sigmų taisyklę — nuo vidurkio trijų sigmų ribose sueina maždaug 99,7% visų x_i dydžių, tai šiuo atveju $\bar{x} \pm 3\sigma_x$ ribose neturėtų būti mažiau kaip 89% visų dydžių. Bet iš kitos pusės, jei dispersija griežtai normali, tai neturėtų sueiti ir visas 100%. Gautąjį $\sigma = 1{,}58$ padauginę trimis rasime, kad $3\sigma = 4{,}74$, taigi:

$$\bar{x} \pm 3\sigma = 5 \pm 3 \,.\, 1{,}58 = 5 \pm 4{,}74 =$$

$$= 0{,}26 \longleftrightarrow 9{,}74.$$

Lentelė 19.

BINOMO BEI JO KOEFICIENTŲ LENTELĖ.

$(a + b)^n = 2^{10}$.

n_i	x_i	$x_i - \bar{x} = \Delta_i$	Δ_i^2	y_i	$\Delta_i^2 y_i$
0	1	2	3	4	5
1	0	—5	25	1	25
2	1	—4	16	10	160
3	2	—3	9	45	405
4	3	—2	4	120	480
5	4	—1	1	210	210
6	5	0	0	252	0
7	6	1	1	210	210
8	7	2	4	120	480
9	8	3	9	45	405
10	9	4	16	10	160
11	10	5	25	1	25
N=11	$\Sigma x_i = 55$	0	$110 = \Sigma\Delta_i^2$	$1024 = \Sigma y_i$	$2560 = \Sigma\Delta_i^2 y_i$

Vidurinis kvadratingasis nukrypimas (σ) pagal formulą

$$\sigma = \sqrt{\frac{\Sigma \Delta_i^2 y_i}{\Sigma y_i}} = \sqrt{\frac{2560}{1024}} = \sqrt{2{,}5}\,(=)\,1{,}58$$

Tuo būdu trijų sigmų nuo vidurkio ribose ideališkai normalios dispersijos st. eilutėj visi dydžiai sueiti negali. Pasižiūrėję į lentelę lengvai pastebėsime, kad jau intervalo ribose tarp 1 ir 9 telpa 1022 dydžiai, t. y. net 99,8%, taigi beveik tiek pat, kaip ideališkai normalios dispersijos taisyklė reikalauja. Nedidelis skirtumas išeina dėl to, kad, kaip buvo minėta, ideališkai normali dispersija yra atstovaujama binominių koeficientų geometriškam pavaizdavime, o ne jų st. lentelėje.

Patsai kvadratinis dispersijos koeficientas sudaro apie 30% vidurkio ir gaunamas pagal formulą XXIV, kur

$$d_\sigma = \frac{1{,}58}{5{,}00} = 0{,}3016 \text{ arba } 30{,}16\%.$$

Jei tuos pačius binomo koeficientus paimsime ne iš $(a + b)^{10}$, o iš $(a + b)^{14}$ prie tos pačios sąlygos, kur $a + b = 2$, o $b = a = 1$, tai gausime šitokį vaizdą:

Lentelė 20.

BINOMO BEI JO KOEFICIENTŲ LENTELĖ, kur $(a+b)^n = 2^{14}$.

n_i	x_i	$x_i - x = \Delta_i$	Δ_i^2	y_i	$\Delta_i^2 y_i$
0	1	2	3	4	5
1	0	—7	49	1	49
2	1	—6	36	14	504
3	2	—5	25	91	2 275
4	3	—4	16	364	5 824
5	4	—3	9	1 001	9 009
6	5	—2	4	2 002	8 008
7	6	—1	1	3 003	3 003
8	7	0	0	3 432	0
9	8	1	1	3 003	3 003
10	9	2	4	2 002	8 008
11	10	3	9	1 001	9 009
12	11	4	16	364	5 824
13	12	5	25	91	2 275
14	13	6	36	14	504
15	14	7	49	1	49
			$\Sigma \Delta_i^2 = 280$	$\Sigma y_i = 16\,384$	$\Sigma \Delta_i^2 y_i = 57\,344$

Vidurinis kvadratingasis nukrypimas σ pagal formulą

$$\sigma = \sqrt{\frac{\Sigma(\Delta_i^2 y_i)}{\Sigma y_i}} = \sqrt{\frac{57344}{16384}} = \sqrt{3{,}5}\,(=)\,1{,}87$$

Dispersijos koeficientas $d_\sigma = \frac{\sigma}{\bar{x}}$ šiuo atveju išeina:

$$d_\sigma = \frac{\sigma}{\bar{x}} = \frac{1{,}87}{7{,}00} = 0{,}2671 \text{ arba } 26{,}71\%.$$

Taigi esmėje visai tos pačios rūšies st. eilutei su ta pačia griežtai normaline dispersija patsai d_σ koeficientas išeina mažesnis, jei skaitomųjų vienetų skaičius (n) yra didesnis (vad. didžiųjų skaičių dėsnio pasireiškimas).

Pagal trijų sigmų taisyklę, trijų kvadratingųjų nukrypimų ribose ir šiuo atveju toks pat skaitomųjų vienetų skaičius sueina (apie 99,7%).

Įrodymas:

kai $\sigma = 1{,}87$, tai $3\sigma = 3 \cdot 1{,}87 = 5{,}61$.

O kai $\bar{x} = 7$, tai

$$\bar{x} \pm 3\sigma = 7 \pm 5{,}61 = 1{,}39 \longleftrightarrow 12{,}61.$$

Iš lentelės galima nustatyti, kad intervale tarp $x_3 - x_{11}$ imtinai yra susikaupę 16 172 eilutės vienetų n. Taigi, jau šitame intervale telpa $\frac{16\,172}{16\,384}$ (=) 0,987 arba apie 98,7% visų dydžių.

Panašiu būdu galima surasti ir kiek sk. vienetų sutelpa vieno ar dviejų kvadratinguųjų nukrypimų ribose.

Kai $\bar{x} = 7$, o $\sigma = 1{,}87$,

tai $\bar{x} \pm 1\sigma = 7 \pm 1{,}87 = 5{,}13 \longleftrightarrow 8{,}87$.

Taip pat

$$\bar{x} \pm 2\sigma = 7 \pm 3{,}74 = 3{,}26 \longleftrightarrow 10{,}74.$$

Šiose ribose ir turi sueiti apytikriai 68,3% event. 95,4% visų sk. vienetų.

Viduriniu kvadratinguoju nukrypimu pasikliaujant surandamas ir asimetrijos koeficientas pagal formulą:

$$a_\sigma = \frac{3(\bar{x} - M)}{\sigma}.$$

Bet kadangi prie griežtai normalios dispersijos, taigi, ir šiuo atveju, aritmetinis vidurkis ir medianas sutampa, t. y. $\bar{x} = M$, tat asimetrijos koeficientas išeis 0. Tačiau, jeigu tik st. eilutės dydžiai pasiskaidę nesimetriškai, kad ir labai nedaug, tai asimetrijos koeficientas išeis kuris nors teigiamas ar neigiamas skaičius. Iš šio skaičiaus galima jau spręsti, ypač ryšyj su dispersijos koeficientu, kaip didelė st. eilutės asimetrija, ir ar pati dispersija dar turi ką nors bendra su normalios dispersijos tipu ar gal jau priklauso kurio nors kito, t. y. nenormalios dispersijos tipo.

Dabar panagrinėkime, ar galima (ar kiek galima) iš griežtai normalios dispersijos dėsnių ir tuos dėsnius nustatančių st.

skaičių ištirti bet kurį socialinio ar ekonominio gyvenimo reiškinį. Pavyzdžiui paimkime tuos pačius kviečių derliaus duomenis, kuriuos naudojome įvairių tipų vidurkiams rasti. Derlius iš 1 ha atskirose apskrityse laikydami x-inės eilutės dydžiais, gausime šitokią kviečių derliaus ir jų dispersijos lentelę:

Kviečių derliaus vidurinis kvadratingasis nukrypimas pagal formulą:

$$\sigma = \sqrt{\frac{\Sigma \triangle_i^2 y_i}{\Sigma y_i}}.$$

gaunamas σ (=) 1,04 q 1 ha, nes:

$$\sigma = \sqrt{\frac{\Sigma \triangle_i^2 y_i}{\Sigma y_i}} = \sqrt{\frac{174{,}4085}{162{,}6}}$$

(=) $\sqrt{1{,}074}$ (=) 1,04.

Kai aritmetinis vidurkis $\bar{x}$ = 12,95 q, tai dispersijos koeficientas pagal formulą XXIV

$$d_\sigma = \frac{\sigma}{\bar{x}} = \frac{1{,}04}{12{,}95} \text{ (=) } 0{,}08 \text{ arba apie } 8\%.$$

Taigi dispersijos koeficientas išeina net žymiai mažesnis negu griežtai normalios dispersijos binomo koeficientų eilutei, kur buvo n = 16 384. Stebėtis netenka, nes šiuo atveju pasėlių ploto hektarų skaičius n = 162 600, taigi beveik 10 kartų didesnis.

Kvadratingasis asimetrijos koeficientas a_σ pagal XXV formulą, kai $\bar{x}$ = 12,95 ir M = 13,00, gaunamas — 0,14, nes:

$$a_\sigma = \frac{3(m - M)}{\sigma} = \frac{3(12{,}95 - 13{,}00)}{1{,}04} =$$

$$= \frac{3 \cdot -0{,}05}{1{,}04} = \frac{-0{,}15}{1{,}04}$$

(=) —0,14.

Kai asimetrijos koeficientas išeina artimesnis 0, ne 1, o dispersijos koeficientas nesiekia 50% aritmetinio vidurkio, tai yra žymė, jog st. eilutės dispersija iš artutinio, t. y. statistikai dar patikimo, normalumo ribų neišėjusi. O jeigu taip, tai ir toli-

KVIEČIŲ DERLIAI 1932—1935 M.

Lentelė 21.

Nr.	Apskritys	Derlius iš 1 ha kvintalais (100 kg.)	Pasėlių plotas 1000 ha	Bendras derlius 1000 kvint.	Atskirose apskrityse derlių nukrypimai nuo vidurkio $\bar{x} = 12{,}95$	Derliaus nukrypimų kvadratai	Derliaus nukrypimų kvadratų sandauga su pasėlių plotu
		x_i	y_i	$x_i y_i$	$\Delta_i = x_i - \bar{x}$	Δ_i^2	$\Delta_i^2 y_i$
	0	1	2	3	4	5	6
1	Alytus	11,90	7,8	92,82	—1,05	1,1025	8,59950
2	Biržai	13,60	12,8	174,08	0,65	0,4225	5,40800
3	Kaunas	11,60	10,8	125,28	—1,35	1,8225	19,68300
4	Kėdainiai	13,30	14,4	191,52	0,35	0,1225	1,76400
5	Kretinga	11,90	2,6	30,94	—1,05	1,1025	2,86650
6	Marijampolė	13,00	10,6	137,80	0,05	0,0025	0,02650
7	Mažeikiai	13,60	7,1	96,56	0,65	0,4225	2,99975
8	Panevėžys	12,70	15,2	193,04	—0,25	0,0625	0,95000
9	Raseiniai	13,40	10,7	143,38	0,45	0,2025	2,16675
10	Rokiškis	13,30	4,0	53,20	0,35	0,1225	0,49000
11	Seinai	11,20	2,0	22,40	—1,75	3,0625	6,12500
12	Šakiai	12,60	8,5	107,10	—0,35	0,1225	1,04125
13	Šiauliai	13,70	24,4	334,28	0,75	0,5625	13,72500
14	Tauragė	13,30	4,5	59,85	0,35	0,1225	0,55125
15	Telšiai	10,50	2,2	23,10	—2,45	6,0025	13,20550
16	Trakai	11,30	2,4	27,12	—1,65	2,7225	6,53400
17	Ukmergė	11,50	6,8	78,20	—1,45	2,1025	14,29700
18	Utena	11,00	4,6	50,60	—1,95	3,8025	17,49150
19	Vilkaviškis	15,10	8,2	123,82	2,15	4,6225	37,90450
20	Zarasai	11,40	1,3	14,82	—1,55	2,4025	3,12325
21	Klaipėda	14,70	1,2	17,64	1,75	3,0625	3,67500
22	Pagėgiai	18,20	0,4	7,28	5,25	27,5625	11,02500
23	Šilutė	15,70	0,1	1,57	2,75	7,5625	0,75625
		298,50 = $= \Sigma x_i$	162,6 = $= \Sigma y_i$	2106,40 = $= \Sigma x_i y_i$		69,0975 = $= \Sigma \Delta_i^2$	174,40850 = $= \Sigma \Delta_i^2 y_i$

mesnis tokios st. eilutės analizis, taip pat ir atitinkamos išvados, yra galimos pasikliaujant tikimybių teorijos dėsniais, galiojančiais atsitiktinių įvykių didiesiems skaičiams.

Visų pirma, tokiais atvejais turi galioti, aplamai imant, vad. trijų sigmų taisyklė. Kai kviečių derliaus vidurkis $\bar{x} = 12{,}95$ q, o vidurinis kvadratingasis nukrypimas $\sigma = 1{,}04$ q, tai pagal trijų sigmų taisyklę:

	turėtų tilpti apie		bet daugiau kaip
$\bar{x} \pm 1\sigma$	apie	68,3% visų derlių	0
$\bar{x} \pm 2\sigma$	„	95,5% „ „	75%
$\bar{x} \pm 3\sigma$	„	99,7% „ „	89,9%

Patikrinę gauname:

	intervalą,	kuriame iš 162 600 ha bendro ploto sutelpa:
$12{,}95 \pm 1{,}04$	11,91 — 13,99	122 600 arba 75,40%
$12{,}95 \pm 2 \cdot 1{,}04$	10,87 — 15,03	151 700 „ 93,30%
$12{,}95 \pm 3 \cdot 1{,}04$	9,83 — 16,07	162 200 „ 99,97%

Taigi trijų kvadratingųjų nukrypimų ribose sutelpa net daugiau, kaip prie griežtai normalios dispersijos. Iš šito pavyzdžio matyti, kad vid. kvadratingasis nukrypimas σ, kaip priemonė st. eilučių dispersijos ypatybėms pažinti, turi didelę statistinę prasmę. Juo labiau, kad statistikos praktikoj dažniausia turima reikalų su aprėžtu skaitomųjų vienetų skaičiumi N, taigi, intervalą praplėtus ligi keturių ar penkių sigmų apie vidurkį, praktiškai galima apimti visą 100% masės vienetų. Tuo tarpu ideališkai normali dispersija, atstovaujama binomo koeficientų, net ir $\bar{x} \pm 100\sigma$ ribose tikrenybės, t. y. 1 arba 100% negarantuoja, nes čia n nėra aprėžti skaičiai, vadinasi, kinta be galo, — nors jau $\bar{x} \pm 4\sigma$ ribose sutelpa 99,9973% visų dydžių.

Kad dar labiau įsisąmonintume vidurinio kvadratingojo nukrypimo, kaip st. eilučių mato, reikšmę, paimkime labai jau asimetriškai pasiskaidžiusią st. eilutę iš Lietuvos žemės ūkio surašymo. Iš to surašymo paskelbtųjų duomenų, kur dydžio klasės turi didelius ir skirtingus intervalus, kvadratingojo nukrypimo suradimas šiek tiek komplikuojasi, nes atskirų dydžio klasių y_i vienetai neturi griežtai aprėžtų x_i dydžių, o tik intervalą (k_i). Bet panašios rūšies kliūtims nugalėti matema-

tinė statistika turi ne vieną priemonę ir ne vieną skaičiavimo metodą. Tačiau šiuo atveju, kai norime kvadratingąjį nukrypimą panaudoti kaip dispersijos matą skirtingose aplinkybėse, paties skaičiavimo metodo nekeisime, o panaudosime tik tas pagalbines priemones, kurias jau naudojome svertinio vidurkio ir kvartilinių dydžių ieškodami. Vadinasi, eilutės x_i dydžiams atstovauti imsime atitinkančias dydžio klases k_i aritmetinį vidurkį, kurį žymėsime $\ddot{x}_i$, o x_i nuo vidurkio nukrypimo gausis jau $\triangle_i = x_i - \bar{x}$. Suradę gi kvadratingąjį nukrypimą σ pagal jau pažįstamą formulą XXVIII, trijų sigmų taisyklei patikrinti reikalingus intervalus, taip pat tuos intervalus atitinkančius ūkių vienetų skaičius y_i, rasime iš esmės tuo pačiu metodu, kaip radome intervalus ir juos atitinkančius vienetų skaičius kvartiliniams dydžiams.

Lentelė *22.*

LIETUVOS ŽEMĖS ŪKIŲ KVADRATINGAJAM NUKRYPIMUI SKAIČIUOTI SCHEMA.

Eilės nr.	Dydžio klasės	Dydžio klasių vidurkiai ($\ddot{x}_i$)	Ūkių skaičius (y_i)	Klasinių vidurkių skirtumai nuo aritmetinio vidurkio ($\bar{x} = 15{,}03$) ($\ddot{x}_i - \bar{x} = \triangle_i$)	Skirtumų kvadratai ($\triangle_i^2$)	Skirtuminių kvadratų sandauga ($\triangle_i^2 y_i$)
0	1	2	3	4	5	
1	1—2	1,37	13 797	—13,66	186,5956	2 574 459
2	2—5	3,41	39 666	—11,62	135,0244	5 355 878
3	5—8	6,44	45 882	— 8,59	73,7881	3 385 546
4	8—10	8,89	32 355	— 6,14	37,6996	1 219 771
5	10—12	10,86	28 060	— 4,17	17,3889	487 933
6	12—15	13,30	31 512	— 1,73	2,9929	94 312
7	15—20	17,15	33 236	2,12	4,4944	149 376
8	20—30	24,17	34 197	9,14	83,5396	2 856 804
9	30—50	37,04	20 597	22,01	484,4401	9 978 013
10	50—100	65,63	6 476	50,60	256,0360	1 658 089
11	100—150	119,38	906	104,35	10888,9225	9 865 364
12	150—200	169,31	379	154,28	23802,3184	9 021 079
13	200 ir d.	318,77	317	303,74	92257,9876	29 245 782
			287 380			75 892 406

Vidurinis kvadratingasis nukrypimas pagal formulą gaunamas:

$$\sigma = \sqrt{\frac{\Sigma \triangle_i^2 y_i}{\Sigma y_i}} =$$

$$= \sqrt{\frac{75\,892\,406}{287\,380}} (=) \sqrt{264{,}1} (=)$$

(=) 16,25 ha.

Kai aritmetinis vidurkis $\bar{x} = 15{,}03$ ha, o medianas $M = 10{,}86$ ha, tai asimetrijos koeficientas pagal formulą:

$$a_\sigma = \frac{3\,(m - M)}{\sigma} =$$

$$= \frac{3 \,.\, (15{,}03 - 10{,}86)}{16{,}25} = \frac{3 \,.\, 4{,}17}{16{,}25} = \frac{12{,}51}{16{,}25} (=) 0{,}77.$$

Palyginę su kvartiliniu dispersijos koeficientu (psl. 202), kuris tesiekė tik 0,25 ha, matome didelį skirtumą.

Dispersijos koeficientas pagal formulą:

$$d_\sigma = \frac{\sigma}{\bar{x}} = \frac{16{,}25}{15{,}03} (=) 1{,}08 \text{ arba beveik } 108\%.$$

O kai asimetrijos ir dispersijos koeficientai artinasi 1, yra aiškus įrodymas, kad st. eilutė yra tolima normalios dispersijos tipui. Be to, vienas vidurkis šiose sąlygose negali charakterizuoti visą masę. Statistikos praktikoj panašiu atveju st. masė skaidoma dydžio klasėmis, jei duomenys gauti ištisinio stebėjimo keliu, ir vienu ar kitu principu koriguojami (papildomais stebėjimais, kraštutiniai dydžiai eliminuojami) — jei duomenys gauti reprezentaciniu metodu.

Prie tos progos įsidėmėtina, kad net ir šiose sąlygose vad. trijų sigmų taisyklė galios nėra nustojusi. Jau vieno kvadratingojo nukrypimo riba apie aritmetinį vidurkį apima visus be išimties smulkesnius ir vidutinius ūkius ligi 31,28 ha imtinai, nes intervalas k_i pagal formulą:

$$k_i = \bar{x} \pm 1\sigma = 15{,}03 \pm 16{,}25 = 1 - 31{,}28 \text{ ha}^7.$$

Tuo tarpu intervale ligi 30 ha sueina 258 705 arba 89,8% visų ūkių. Imant gi intervalą ligi 31,28 ha, gaunama $Sy_i = 259\,014$ ūkių, kurie sudaro visa 90% ūkių.

Ieškant kiekvienam $k_{\bar{x}} \pm n\sigma$ intervalui apytikrio ūkių skaičiaus n_i galima pasikliauti šia formula:

$$n_i = \left[\frac{(\bar{x} + 1\sigma) - k_{i,1}}{k_{i,2} - k_{i,1}} \cdot y_i + Sy_i\right] - \left[\frac{(\bar{x} - 1\sigma) - k_{-i,1}}{k_{-i,2} - k_{-i,1}} \cdot y_i + Sy_{-i}\right];$$

Šioj formuloj $k_{i,1}$ ir $k_{-i,1}$ atstovauja atitinkamų klasių intervalo žemutinę (pradinę) ribą, $k_{i,2}$ ir $k_{-i,2}$ tos pat klasės intervalo aukštutinę (paskutinę) ribą, y_i ir y_{-i} yra tos pačios klasės dydžių skaičiai, o Sy_i ir Sy_{-i} atstovauja visų x_i dydžių, bendrąsias sumas y_i ligi priešpaskutinių dydžio klasių k_i.

Mūsų paimtam pavyzdžiui antroji formulos dalis atkrinta, nes $\bar{x} - 1\sigma$ sudaro iš viso galimą žemutinio intervalo ribą $k_i = 0$. Jei sutarsime atitinkamo sigmų skaičiaus apibrėžiamus intervalus ir juos atitinkančius skaičius žymėti skaičiais, tai gausime:

$$n_1 = Sy_8 + \frac{[(\bar{x} + 1\sigma) - k_{i,1}] \cdot y_9}{k_{i,2} - k_{i,1}} =$$

$$= \frac{(15{,}03 + 16{,}25) - 30}{50 - 30} \cdot 20\,597 + 258\,705 =$$

$$= \frac{31{,}28 - 30{,}00}{20} \cdot 20\,597 + 258\,705 =$$

$$= \frac{1{,}28 \cdot 20\,597}{20} + 258\,705 =$$

$$(=) \frac{26\,364}{20} + 258\,705\ (=)\ 1\,318 + 258\,705 =$$

$$(=)\ 260\,023 \text{ ūkiai.}$$

---

[7] Žemutinė intervalo riba nepaduodama, nes $\bar{x} - 1\sigma = 15{,}03 - 16{,}25 = -1{,}22$ ha. O tokio ūgio būti negali.

Kai bendrasis ūkių skaičius N = 287 380, o $\bar{x} \pm 1\sigma$ ribose telpa $n_1 = 260\,023$, tai $\frac{n_1}{N} = \frac{260\,023}{287\,380}$ (=) 0,9 arba 90% visų ūkių. O tai yra žymiai daugiau, negu normalios dispersijos pirmosios sigmos norma (68,3%). Panašiu būdu skaičiuodami n_2 gauname:

$$n_2 = \frac{(\bar{x} + 2\sigma) - k_{2,1}}{k_{2,2} - k_{2,1}} \cdot y_9 + Sy_8 =$$

$$= \frac{(15{,}03 + 2 \cdot 16{,}25) - 30}{50 - 30} \cdot 20\,597 + 258\,705 =$$

$$= \frac{17{,}53 \cdot 20\,597}{20} + 258\,705 =$$

$$= 276\,758 \text{ ūkiai.}$$

Tuo būdu

$$\frac{n_2}{N} = \frac{276\,758}{287\,380} \text{ (=) } 0{,}963 \text{ arba } 96{,}3\%.$$

Taip pat ir dėl $m \pm 3\sigma$ gauname n_3

$$n_3 = \frac{(\bar{x} + 3\sigma) - k_{3,1}}{k_{3,1} - k_{3,1}} \cdot y_{10} + Sy_9 =$$

$$= \frac{(15{,}03 + 3 \cdot 16{,}25) - 50}{100 - 50} \cdot 6\,476 + 279\,302 =$$

$$= \frac{63{,}78 - 50}{50} \cdot 6\,476 + 279\,302 =$$

$$= \frac{13{,}78 \cdot 6\,476}{50} + 279\,302 = 1\,784 + 279\,302 =$$

$$= 281\,086 \text{ ūkiai.}$$

Taigi trijų sigmų ribose apie vidurkį telpa:

$$\frac{n_3}{N} = \frac{281\,086}{287\,380} \text{ (=) } 0{,}98 \text{ arba } 98\%.$$

Tuo būdu matome, kad net ir gana nesimetriškai pasiskaidžiusioj st. eilutėj vieno kvadratingojo nukrypimo ribose nuo

vidurkio sueina apie 90% visų dydžių, dviejų kvadratinguųjų nukrypimų ribose beveik 96% (beveik norma) ir trijų — 98%.

Taigi tikimybė, kad iš visos masės bet kuris atsitiktinai paimtas ūkis yra

$$\bar{x} \pm 3\sigma = 15{,}03 \pm 3 \,.\, 16{,}25 = 1 \text{ —— } 63{,}78 \text{ ha ribose,}$$

yra lygi 98%, ir tik 2% tikimybės už tai, kad gali būti pataikytas ūkis, kuris yra didesnis kaip 63,78 ha.

e) **Apytikriai skaičiuojamas metodas kvadratingiems nukrypimams surasti.**

Kvadratingieji nukrypimai gali būti nustatomi betarpiškai ir iš x-inės eilutės dydžių, neieškant $x_i - \bar{x}$ skirtumų. Šiuo atveju, tačiau, x_i dydžių kvadratų suma (Σx_i^2), prieš traukiant kvadratinę šaknį, reikia sumažinti aritmetiniu vidurkiu kvadrate $(\bar{x}^2)$ pagal formulą:

$$\sigma = \sqrt{\frac{\Sigma x_i^2}{n} - \bar{x}^2} \qquad \text{(XXVIIa).}$$

$$\sigma = \sqrt{\frac{\Sigma x_i^2 y_i}{\Sigma y_i} - \bar{x}^2} \qquad \text{(XXVIIb).}$$

Lentelė 23.

XXVII FORMULOS A IR B VARIJANTŲ PRITAIKOMOJO SKAIČIAVIMO SCHEMA.

Nr.	Varijantas a		Varijantas b		
	(x_i)	(x_i^2)	(y_i)	$(x_i^2 y_i)$	$(y_i x_i)$
0	1	2	3	4	5
1	2	4	2	8	4
2	3	9	3	27	9
3	4	16	4	64	16
4	5	25	3	75	15
5	6	36	2	72	12
	$\Sigma x_i = 20$	$\Sigma x_i^2 = 90$	$\Sigma y_i = 14$	$\Sigma y_i x_i^2 = 246$	$\Sigma y_i x_i = 56$

Duotosioms skaičių eilutėms turime aritmetinį vidurkį $\bar{x} = 4$, x-inės eilutės dydžio kvadratų suma $\Sigma x_i^2 = 90$. Taigi vid. kvadratingasis nukrypimas pagal formulą XXIIa:

$$\sigma = \sqrt{\frac{\Sigma x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{90}{5} - 4^2}$$

$$= \sqrt{18 - 16} = \sqrt{2} \ (=) \ 1{,}41.$$

Formulai XXVIIb gauname:

$$\bar{x} = \frac{\Sigma y_i x_i}{\Sigma y_i} = \frac{56}{14} = 4, \ \Sigma y_i = 14 \text{ ir } \Sigma y_i x_i^2 = 246.$$

Taigi

$$\sigma = \sqrt{\frac{\Sigma x_i^2 y_i}{\Sigma y_i} - \bar{x}^2} = \sqrt{\frac{246}{14} - 4^2}$$

$$(=) \sqrt{17{,}8 - 16} = \sqrt{1{,}8} \ (=) \ 1{,}34.$$

Tačiau kvadratingojo nukrypimo suradimas betarpiškai iš originalinių st. eilutės dydžių tik tuo atveju gali sutaupyti laiko, kai patys x_i dydžiai, ar jų gausumo koeficientai (y_i), yra maži. Jei gi x_i dydžiai ar jų gausumą atstovaujantys skaičiai (y_i) dideli (iš keleto skaitmenų), tai darbas nesumažėja.

Kai x_i ar y_i eilučių skaičiai dideli, arba st. duomenys yra jau suskirstyti į dydžio grupes vienodais intervalais, galima tenkintis artutinio skaičiavimo metodais panašiai kaip vidurkiams. Vadinasi, reikia sudaryti dydžio grupes arba klases (k) vienodais intervalais ir x-inės eilutės y_i grupių varijantams duoti kiekvienos grupės intervalo vidurinį dydį. Sutarę šį dydį žymėti x', jį surasime pagal formulą $x' = \frac{k_2 + k_1}{2}$. Ir šiuo būdu st. eilutėmis x' dydžių kuris nors vienas parenkamas konstanta (a) ir ieškomas vidurinis kvadratingasis nukrypimas iš x' skirtumų nuo konstantos a panašiai, kaip ieškoma vidurkių pagal formulą VIIIc. Jei sutarsime konstantai paimti $\dot{x} = x'_i - x'_0$ bus teigiami arba neigiami dydžiai sulig tuo, ar jie už x'_0 didesni ar mažesni. Tuo būdu kvadratingajam nukrypimui formulą galėsime rašyti šitaip:

$$\sigma = \sqrt{\frac{\Sigma y_i \ddot{x}^2}{\Sigma y_i} - \left(\frac{\Sigma \ddot{x} y_i}{\Sigma y_i}\right)^2} \qquad \text{(XXVIIc)}.$$

Pavyzdiniam skaičiavimui paimkime tuos pačius kviečių derliaus duomenis, kurie buvo naudojami vidurkiams skaičiuoti. Iš lentelės 21 galime įsitikinti, kad 23 apskrityse derliai iš 1 ha svyravo tarp 10,5 ir 18,2 q (kvintalų). Suskirstę tuos pačius duomenis į dydžio klases (k) vienodais intervalais, kur k = 1 q (kvintalas arba 100 kg), gauname 9 dydžio grupes, kur visi skirtuminiai dydžiai ($\ddot{x} = x'_i - x'_0$) gauna atitinkamus mato vienetais pareikštus skaičius y_i (pasėlių plotai 1000 ha).

Lentelė 24.

1932—34 MT. KVIEČIŲ DERLIAI LIETUVOJE IR JŲ KVADRATINGIEJI NUKRYPIMAI.

Eil. nr. n	Derliai iš 1 ha apvalintais q (x'_i)	Derlių skirtumai sulig konstanta ($\ddot{x}=x'_1-x'_0$)	Derliaus skirtumų kvadratai ($\ddot{x}^2$)	Derliaus x-inių grupių plotai 1000 ha (y_i)	Derliaus skirtumų ir plotų sandauga ($y_i\ddot{x}$)	Derliaus skirtumo kvadratų ir plotų sandauga ($y_i\ddot{x}^2$)
0	1	2	3	4	5	6
1	10	—3	9	2,2	— 6,6	19,8
2	11	—2	4	38,3	—76,6	153,2
3	12	—1	1	23,7	—23,7	23,7
4	13	0	0	88,5	0,0	0,0
5	14	1	1	1,2	1,2	1,2
6	15	2	4	8,3	16,6	33,2
7	16	3	9	0,0	0,0	0,0
8	17	4	16	0,0	0,0	0,0
9	18	5	25	0,4	2,0	10,0
		$\Sigma\ddot{x}=9$	$\Sigma\ddot{x}^2=69$	$\Sigma y_i=162,6$	$\Sigma y_i\ddot{x}=$ $=-87,1$	$\Sigma y_i\ddot{x}^2=$ $=241,1$

Patiektoj lentelėj turime visus reikalingus skaičius. Kai $\Sigma\, y_i\ddot{x} = -87,1$, $\Sigma y_i\ddot{x}^2 = 241,1$ ir $\Sigma y_i = 162,6$, tai ir kvadratingasis nukrypimas pagal formulą XXVIIc gaunamas:

$$\sigma = \sqrt{\frac{\Sigma y_i \ddot{x}^2}{\Sigma y_i} - \left(\frac{\Sigma y_i \ddot{x}}{\Sigma y_i}\right)^2}$$

$$= \sqrt{\frac{241,1}{162,6} - \left(\frac{-87,1}{162,6}\right)^2} \; (=) \sqrt{1,48 - 0,29} =$$

$$= \sqrt{1,19} \; (=) \; 1,09.$$

Tuo būdu apytikris vid. kvadratingasis nukrypimas σ (=) 1,09 q. Palyginę šitą apytikrį σ su griežtai išskaičiuotuoju pagal duomenis lentelėj 21, randame skirtumą (1,09 — 1,04), kuris nesudaro nė 5%.

Iš tų pačių duomenų išskaičiuotas aritmetinis vidurkis pagal VIIIc formulą beveik sutampa. Nes kai $a = \frac{13 + 14}{2} =$ = 13,5, tai

$$\bar{x} = a + \frac{\Sigma y \ddot{x}}{\Sigma y} = 13,5 + \frac{-87,1}{162,6} \; (=) \; 13,50 - 0,54 = 12,96 \text{ q}$$

prieš 12,95 q gautus iš betarpiško skaičiavimo.

Sekantis skaičiavimo pavyzdys parodo, kad artutinių st. dydžių patikimumas nedaug sumažėja, jei jie skaičiuojami ir iš žymiai mažesnio dydžio grupių skaičiaus, t. y. iš dydžio grupių su didesniais intervalais.

Lentelė 25.

1932—34 MT. KVIEČIŲ DERLIAI LIETUVOJE IR JŲ KVADRATINGIEJI NUKRYPIMAI.

(n) Eilės nr.	Derliai grupėms po 2 q iš 1 ha (k_j)	Derliai iš 1 ha apvalintais q ($x' = \frac{k_2 + k_1}{2}$)	Derlių skirtumai sulig konstanta $x'_o = 13$ ($\ddot{x} = x_j - x_o$)	Derliaus skirtumų kvadratai ($\ddot{x}^2$)	Derliaus x-inių grupių plotai 1000 ha (y_i)	Derliaus skirtumų ir plotų sandauga ($y_i\ddot{x}$)	Derliaus skirtumo kvadr. ir plotų sandauga ($y_i\ddot{x}^2$)
0	1	2	3	4	5	6	7
1	10—12	11	— 2	4	40,5	— 81,0	162,0
2	12—14	13	0	0	112,2	0,0	0,0
3	14—16	15	2	4	9,5	19,0	38,0
4	16—18	17	4	16	0,4	1,6	6,4
	—	—	—	$\Sigma \ddot{x}^2 =$ =24	$\Sigma y_i =$ =162,6	$\Sigma y_i \ddot{x} =$ =—60,4	$\Sigma y_i \ddot{x}^2 =$ =206,4

Iš patiektosios lentelės 25 reikalingus skaičius paėmę XXVIIc formulai gauname:

$$\sigma = \sqrt{\frac{206,4}{162,6} - \left(\frac{-60,4}{162,6}\right)^2} \quad (=) \sqrt{1,27 - 0,14} =$$

$$= \sqrt{1,13} \,(=)\, 1,06 \text{ q.}$$

Aritmetinis vidurkis pagal tuos pačius duomenis išeina 12,67 q, nes: konstanta a iš $\frac{12-14}{2} = 13$, taigi

$$\bar{x} = 13,00 + \left(\frac{-60,4}{162,6}\right) \,(=)\, 13,00 - 0,37 = 12,67.$$

Tuo būdu matome, kad artutinio skaičiavimo metodas iš didesnių dydžio grupių davė st. skaičius, kurie yra visai artimi tikrenybei (skirtumas tarp σ_X ir σ_X' tesudaro tik 1,06 — 1,04 = = 0,02, t. y. nepilnus 2%).

Lentelė 26.

KVADRATINGIEJI NUKRYPIMAI IŠ NUOTOLINIŲ EILUTĖS NR. (x).

Eilės nr.	Derliai derėjimo grupėms po 2 q (k_i)	Derliai iš 1 ha apvalintais q ($x'_i = \frac{k_2+k_1}{2}$)	Nuotolin. eil. nr. (x) nuo konstantos $x'_o = 13$ ($\dot{x}$)	Nuotolinių skaičių (x) kvadratai (x^2)	Derliaus x-inių grupių plotai 1000 ha (y_i)	Derliaus plotų sandauga su x ($y_i x$)	Derliaus plotų sandauga su x^2 ($y_i x^2$)
0	1	2	3	4	5	6	7
1	10—12	11	— 1	1	40,5	— 40,5	40,5
2	12—14	13	0	0	112,2	0,0	0,0
3	14—16	15	1	1	9,5	9,5	9,5
4	16—18	17	2	4	0,4	0,8	1,6
	—	—	—	—	$\Sigma y_i =$ $= 162,6$	$\Sigma y_i x =$ $= -30,2$	$\Sigma y_i x^2 =$ $= 51,6$

Kai dydžio grupių intervalai k_i didesni, o pačių grupių skaičius nemažas, kvadratingojo nukrypimo skaičiavimas gali būti daromas ne iš faktinųjų skirtumų tarp $x'_i - x'_o$, o iš nuotolinių (atokumo) $x_i - x_o$ eilutės nr. (x). Šiuo atveju, tačiau, XXVIIc formulos pirmąjį narį reikia padauginti skaičiumi, kuris gaunamas intervalą (k) pakėlus laipsniu (k^2), o antrąjį

narį padauginti tiesiog intervalo skaičiumi k. Atitinkamai pakeitę formulą XXVII gauname:

$$\sigma = \sqrt{\frac{k^2 \Sigma y_i x^2}{\Sigma y_i} - \left(\frac{k \Sigma y_i x}{\Sigma y_i}\right)^2} \qquad \text{(XXVIIc)}.$$

Lentelėj 26 atitinkamai sudoroti tie patys duomenys rodo, kad $\Sigma y_i x = -30{,}2$ ir $\Sigma y_i x^2 = 51{,}6$. Kai $\ddot{x}_i$ grupių intervalas $k = 2$ (nes visoj eilutėj $k_2 - k_1 = 2$), tai kvadratingasis nukrypimas pagal formulą XXVIId išeina:

$$\sigma = \sqrt{\frac{k^2 \Sigma y_i x^2}{\Sigma y_i} - \left[\frac{k \Sigma y_i x}{\Sigma y_i}\right]^2} = \sqrt{\frac{2^2 \cdot 51{,}6}{162{,}6} - \left(\frac{2 \cdot -30{,}2}{162{,}6}\right)^2} =$$

$$= \sqrt{\frac{4 \cdot 51{,}6}{162{,}6} - \left[\frac{-60{,}4}{162{,}6}\right]^2} \; (=) \; \sqrt{1{,}27 - 0{,}14} =$$

$$= \sqrt{1{,}13} \; (=) \; 1{,}06 \text{ q}.$$

Taigi matome, kad ir šiuo atveju gaunami tie patys rezultatai, nors x-iniai skirtumai, taigi visi daugikliai ir st. eilutės dalinių sandauginės sumos, žymiai sumažėja.

Bet dar ir šis metodas duoda daug sunkaus darbo, jei st. eilutė yra labai ilga ir tuo būdu konstantai pasirinktojo dydžio x_0 nuotoliniai skaičiai x pasiekia dviskaičių dydžių. Tokiais atvejais, kaip ir aritmetiniams vidurkiams, patogiausias yra sumavimų metodas.

S u m a v i m ų m e t o d a s vad. k v a d r a t i n g a j a m n u k r y p i m u i surasti yra betarpiškai susietas su tuo pačiu metodu aritmetiniam vidurkiui pagal jau pažįstamą formulą VIIIc, kur:

$$\bar{x} = a + \frac{S_1}{N} = a + \frac{s_1 - s_{-1}}{\Sigma y_i}, \text{ arba}$$

$$= a + \frac{k\,(S_1)}{N} = a + \frac{k\,(s_1 - s_{-1})}{\Sigma y_i}.$$

Kvadratingajam nukrypimui skaičiuoti reiškinys Sy_1 taip pat dalyvauja, nes, kaip toliau pamatysime, formulos XXVIIId reiškinys $\Sigma y\ddot{x} = Sy_1 = S_1$. Lieka dar gauti skaičių, atitinkantį tos pačios formulos pirmąją dalį ($\Sigma y\ddot{x}$). Šį skaičių su-

mavimų metodo formuloj atitinka Sy_2, taigi $\Sigma y\ddot{x}^2 = Sy_2 = S_2$. Ši antrojo laipsnio kumuliatyvinė suma S_2 gaunama pagal formulą:

$$S_2 = s_{-1} + s_1 + 2s_{-2} + 2s_2.$$

Aplamai, kvadratingajam nukrypimui formulos dalomasis reiškinys yra

$$\Sigma (x_i - \bar{x})^2 = S_2 - \frac{S_1^2}{n}.$$

O kadangi $\Sigma(x_1 - \bar{x})^2 = \Sigma\triangle_i^2$ ir $\sqrt{\frac{\Sigma\triangle_i^2}{n}} = \sigma$, tai ir sumavimų metode kvadratingojo nukrypimo formulą galime rašyti:

$$\sqrt{\frac{S_2 - \frac{1}{n} S_1^2}{n}} = \sqrt{\frac{1}{N}\left(S_2 - \frac{S_1^2}{N}\right)} =$$

$$= \sqrt{\frac{1}{\Sigma y_i}\left(Sy_2 - \frac{Sy_1^2}{\Sigma y_i}\right)} \qquad \text{(XXVIId).}$$

Lentelė 27.

SCHEMATINĖ LENTELĖ VIDURINIAM KVADRATINGAJAM NUKRYPIMUI SUMAVIMŲ METODU SURASTI.

(Duomenys kviečių derliaus Lietuvoje)

n	k	x	y_i	$S_o = s_{-o} + y_o + s_o$	$S_1 = s_{-1} - s_i$	$S_2 = s_{-1} + s_1 + 2s_{-2} + 2s_2$
0	1	2	3	4	5	6
1	10—11	—3	2,2	2,2	2,2	2,2
2	11—12	—2	38,3	40,5	42,7	s_{-2}=44,9
3	12 - 13	—1	23,7	s_{-o}=64,2	s_{-1}=106,9	—
4	13—14	0	88,5	y_o=88,5	—	—
5	14—15	1	1,2	s_o= 9,9	s_1= 19,8	—
6	15—16	2	8,3	8,7	9,9	s_2=12,3
7	16—17	3	0,0	0,4	1,2	2,4
8	17—18	4	0,0	0,4	0,8	1,2
9	18—19	5	0,4	0,4	0,4	0,4

Lentelėj 27 tie patys kviečių derliaus duomenys sudoroti sumavimų metodu. Iš lentelės matyti, kad s_{-2} ir s_2 yra gautos kumuliuojant pirmojo laipsnio sumas s_{-1} ir s_1 iš galų į vidurį ligi priešpaskutinės eilutės.

Tuo būdu gauta:

$s_{-1} = 106{,}9$, $s_1 = 19{,}8$, taigi $S_1 = s_{-1} - s_2 = 106{,}9 - 19{,}8 = 87{,}1$.

Taip pat $s_{-2} = 44{,}9$ ir $s_2 = 12{,}3$, taigi:

$$\begin{aligned} S_2 &= s_{-1} + s_1 + 2s_{-2} + 2s_2 \\ &= 106{,}9 + 19{,}8 + 2 \cdot 44{,}9 + 2 \cdot 12{,}3 \\ &= 106{,}9 + 19{,}8 + 89{,}8 + 24{,}6 \\ &= 241{,}1. \end{aligned}$$

Kadangi $N = \Sigma y_i = 162{,}6$, tai kvadratingasis nukrypimas pagal formulą XXVIId išeina:

$$\sigma = \sqrt{\frac{1}{N}\left(S_2 - \frac{S_i}{N}\right)} = \sqrt{\frac{1}{162{,}6}\left(241{,}1 - \frac{87{,}1^2}{162{,}6}\right)}$$

$$= \sqrt{\frac{241{,}1 - 46{,}9}{162{,}6}} = \sqrt{\frac{194{,}2}{162{,}6}}$$

$$(=)\ \sqrt{1{,}19}\ (=)\ 1{,}09.$$

Tuo būdu sumavimų metodas pagal formulą XXVIId duoda visai tuos pačius rezultatus kaip ir žymiai komplikuotesnis artutinio skaičiavimo metodas pagal formulą XXVIIc. O jeigu tarti, kad lentelėse 25 ir 26 panaudoti x-iniai dydžiai yra iš betarpiško stebėjimo gautieji originaliniai skaičiai, tai sumavimų metodu gautąjį kvadratingąjį nukrypimą atitiks ir kvadr. nukrypimą, gautąjį griežtai pasikliaujant diferenciniu metodu pagal kitus XXVII formulos varijantus, kur

$$\sigma = \sqrt{\frac{\Sigma \triangle^2 y_i}{\Sigma y_i}} = \sqrt{\frac{\Sigma x_i^2 y_i}{\Sigma y_i} - \bar{x}^2}.$$

Nes jei tarsime, kad x-inės eilutės dydžiai x_i st. eilutėj prasideda $x_1 = 10$, $x_2 = 11$ ir t. t. ligi $x_9 = 18$, o jų pasikartojimo skaičiai y_i yra tie patys, kaip lentelėj 24 (vadinasi, $y_1 = 2{,}2$, $y_2 = 38{,}3$ ir t. t. ligi paskutiniojo nario $y_9 = 0{,}4$),

tai aritmetinis vidurkis $\bar{x} = 12{,}464$, o patsai v. kvadratingasis nukrypimas, vis viena pagal kurią formulą skaičiuojamas, išeis $\sigma\,(=)\,1{,}1$. Tai visai atitinka σ ir iš sumavimų metodo, jei σ griežtumą aprėžti vienu skaitmeniu po kablelio. Kvadratingasis nukrypimas išskaičiuotas pagal formulą XXVIIb, davė šitokius skaičius:

$\Sigma\bar{x}^2 y_i\,(=)\,156{,}56$, $\bar{x}^2\,(=)\,155{,}35$, taigi

$$\sigma = \sqrt{156{,}56 - 155{,}35} = \sqrt{1{,}21}\,(=)\,1{,}1.$$

Mažas nesutapimas pareina grynai nuo trumpinamojo artutinio skaičiavimo.

Aplamai, sumavimų metodo pranašumas pilnumoj pasireiškia tik pakankamai ilgose st. eilutėse, ir jo pirmenybė prieš kitus metodus bus juo didesnė, juo st. eilutė ilgesnė. Ir atvirkščiai: trumpose, nedaug narių teturinčiose, st. eilutėse sumavimų metodas savo statistinės prasmės nustoja.

Ši sumavimų metodo ypatybė daro jį ypatingai svarbia matematinio analizio priemone įvairioms statistikos įstaigoms, kurioms tenka apdirbinėti bei klasifikuoti iš st. stebėjimo gautąją medžiagą. Betarpiškai arba diferenciniu metodu kvadratingojo nukrypimo ieškojimas, kur įvairaus didumo skaitomųjų vienetų būna ir tūkstančių tūkstančiai, nėra įmanoma. Tuo tarpu sumavimų metodas nedaug laiko užima, be to, jis gali būti naudojamas paraleliškai su kiekybine klasifikacija. Nes bendroji sk. vienetų suma N, kuri sumavimų metode žymima S_0, vistiek reikalinga žinoti. O turint s_{-0} ir s_0 yra lengva gauti s_{-1} ir s_1. Šios gi sumos duoda ne tik S_1, bet, kaip matėme, betarpiškai veda ir prie s_{-2} su s_2, taigi ir prie S_2. Dėl šios priežasties net artutinio skaičiavimo metodai negali lygintis su sumavimų metodu tais atvejais, kai pirmykštė st. medžiaga yra dar neklasifikuota.

Bet šioj vietoj mes jau susiduriame ir su kitu statistinės praktikos klausimu, būtent: iš ko ir kaip pažinti, ar artutinio skaičiavimo metodais gautieji statistikos dydžiai (vidurkiai, kvadratingieji nukrypimai, dispersijos bei asimetrijos koeficientai) tikrenybę atitinka mažiau ar daugiau? Kiek tokiais artutinio skaičiavimo metodais galima pasitikėti? — Atsakymą į

šiuos klausimus tegali duoti tik tikimybių teorija bei tikimybiniai statistikos dydžiai.

Literatūra ir papildomos žinios apie santykinius st. skaičius, vidurkius ir dispersiją yra visuose statistikos vadovėliuose. Atskirai pažymėtini: 1) Dr. W. Winkler, Die statistischen Verhältniszahlen, Leipzig-Wien, 1923, 178 p., 2) Fr. Žižek, Die statistischen Mittelwerte, Leipzig, 1908, 3) Rietz-Bauer, Handbuch der mathemat. Statistik, pp. 29—46 ir 109—158, 4) O. N. Anderson, Einführung in die mathematische Statistik, pp. 128—266, 5) L. March, Principles de la Methode Statistique, pp. 201—415 ir 503—556, 6) A. Bowley, Elements of Statistics, pp. 82—124, 7) V. I. Romanovskij, Elementarnyj kurs matematičeskoj statistiki, pp. 151—223, ir kt.

C. Statistikos dėsningumai ir tikimybės.

7. Statistinės ir matematinės tikimybės.

a) Vad. statistikos dėsningumai ir kai kurie jų pasireiškimai.

Jau skyriuje apie statistikos eilučių dispersiją buvo patiekta įrodymų, kad binomo koeficientų eilutės pasiskaidymas ir kviečių derėjimų iš 1 ha pasiskaidymas atskirose apskrityse turi daug panašumo bei duodasi maždaug vienodais mato dydžiais matuojami. O binomo koeficientų ir derėjimų eilutės yra visai skirtingos kilmės: pirmoji yra grynai teorinių arba aprioristinių matematikos skaičiavimų padarinys, antroji — derlių statistinio, t. y. empirinio stebėjimo vaisius. Vadinasi, ir gyvenimo reiškiniai, bent jau tam tikrais atvejais, vyksta taip pat dėsningai, kaip ir bet kuri matematikos teorijos apibrėžiama tikimybė.

Pavyzdžiui imkime tokius žmonių gyvenimo reiškinius, k. a. mirimai ir vedybos. Nors visi žino, kad vieną kartą reikės mirti, bet, išskiriant vieną antrą desperatą arba vilties nustojusį žmogų, niekas mirties nepageidauja. Atvirkščiai: visomis galimomis patyrimo ir mokslo teikiamomis priemonėmis kiekvienas stengiasi mirtį nuo savęs kiek galint atitolinti. Ir, išskiriant retus savižudybės atsitikimus, galima sakyti, kad jei kurią dieną vienas ar kitas žmogus numiršta, tai toks faktas atitinka įvykių logikos eigą ir be mirštančiojo noro ar jo valios prisidėjimo.

Iš kitos pusės imkime vedybas. Čia, rodos, vedasi tik tie, kurie nori, ir tik tada, kada nori. O jei kas vesti nenori — palieka nevedęs. Tegu kartais tėvai, giminės arba draugai savo patarimais bei sugestijomis ir prisideda prie vedybų pagreitinimo, pavėlinimo ar atidėjimo. Bet, aplamai, vedybos vyksta visuomet pačių apsivedančiųjų valia ir jų pritarimu, ne taip kaip mirimai. Taigi atrodytų, kad vedybų ir mirimų atsitikimai atskirais laiko periodais taip pat turėtų būti skirtingi.

Bet štai paimkime Lietuvos vedybų ir mirimų duomenis.

Lentelė 28.

LIETUVOJE 1000-ČIUI GYVENTOJŲ TENKA:

	vedybų	mirimų
1921	7,4	15,2
1922	7,7	17,7
1923	7,1	15,1
1924	8,1	16,3
1925	7,9	16,9
1926	8,5	15,4
1927	8,6	17,3
1928	8,3	15,6
1929	7,5	17,1
1930	7,8	15,9
1931	8,0	15,8
1932	7,9	15,3
1933	8,1	13,5
1934	7,4	14,6
1935	7,4	14,0
1936	7,5	13,4
1937	7,4	13,2

Pasižiūrėjus čia patiektiems skaičiams galima paklausti: kas vertė Lietuvos gyventojus per ištisus 16 metų kas metai iš kiekvienų 1000 ne mažiau kaip 7,1 ir ne daugiau kaip 8,6 apsivesti? Ir dėlko per tą patį laiką prieš Lietuvos gyventojų norą vis dėlto kas metai iš tų pačių 1000 gyventojų pasimirė ne mažiau kaip 13,2, bet ne daugiau kaip 17,7 žmonių?

Kaip matome, įvykiai, kurie atrodo pareiną išimtinai nuo žmonių valios (vedybos) pasireiškia tiek pat tvarkingai ir dėsningai, kaip ir įvykiai, kurie atsitinka prieš žmonių valią ir norą. Bet štai paimkime tuos pačius vedybų ir mirimų duomenis ir suskaidykime smulkesniais laiko periodais, pav. mėnesiais.

Lentelė 29a.

MĖNESINIAI JUNGTUVIŲ SKAIČIAI.

Eil. Nr.	Mėnesiai	1935	1934	1933	1933—1935
0	1	2	3	4	5
1	Sausis . . .	1 290	1 967	1 668	1 641
2	Vasaris . . .	3 078	4 596	6 437	4 704
3	Kovas . . .	2 921	304	305	1 177
4	Balandis. . .	900	1 369	818	1 029
5	Gegužis . . .	1 040	1 596	1 254	1 297
6	Birželis . . .	1 562	1 087	1 471	1 373
7	Liepa . . .	1 144	1 003	1 248	1 132
8	Rugpiūtis . .	1 161	920	943	1 008
9	Rugsėjis. . .	1 071	1 142	1 042	1 085
10	Spalis . . .	1 180	1 230	1 311	1 240
11	Lapkritis . .	2 058	2 041	2 038	2 046
12	Gruodis. . .	849	991	976	939
		18 254	18 246	19 511	18 671

Patiektieji duomenys jau rodo, kad, suskaidžius smulkesniais skaitomaisiais periodais, vedybų ir mirimų svyravimai yra didesni. Tačiau kai mirimų svyravimai, palyginti, vis dar nedideli, tai vedybų svyravimai atskirais mėnesiais net labai dideli. Pav., tų pačių 1934 m. vasario mėn. apsivedė 4596 poros, o kovo mėn. tik 304 poros, t. y. — 15 kartų mažiau. Panašus reiškinys pastebimas vasario ir kovo mėn. vedybose ir 1933 m. Tačiau 1935 m. kovo mėn. būta beveik tiek pat vedybų kaip ir vasario mėn. Taip pat žymiai išsiskiria vedybų skaičiai randami kai kada ir balandžio mėn., be to, visuomet mažiau vedybų būna gruodžio mėn. Ką visa tai reiškia?

Dėl nepaprastai didelių svyravimų atskirais mėnesiais vedybų atsitikimuose visų pirma tenka prisiminti pas mus galiojantieji kai kurie papročiai. Pav., yra žinoma, kad katalikų

bažnyčia nelabai proteguoja vedybas gavėnios ir adventų metu. Ir tikrai mes randame gruodžio mėn. (adventas) vedybų būna tik apie pusę tiek, kiek lapkričio ar sausio mėn. O 1935 m., kai Velykos buvo vėlybos, mažiausia vedybų randama balandžio mėn. Taigi bažnyčios arba tikybinių įsitikinimų sugestija tam tikrais metų tarpais susilaikyti nuo vedybų pasireiškia nepaprastais vedybų susikaupimais prieš tuos periodus. Pav., 1934 m. paskutinį mėnesį prieš gavėnią buvo net 6437 vedybos, t. y. 4 kartus daugiau kaip vidutiniškai kas mėnuo ir, be to, paskutinį mėnesį prieš adventus 2038 vedybų, t. y. 47% daugiau, kaip vidutiniškai kas mėnuo. Tuo būdu atrodo, lyg būtų ištisiems metams iš anksto paskirtas tam tikras vedybų skaičius; ir jei kurį nors laiko tarpą dėl kurių nors priežasčių vedybų įvykti negali, tai kandidatai jau prieš terminą pasiskubina apsivesti, kad tik nepasivėlintų ir bendrojo metinio vedybų skaičiaus nesumažintų.

Lentelė 29b.

MĖNESINIAI MIRIMŲ SKAIČIAI.

Eil. Nr.	Mėnesiai	1935	1934	1933	1933—1935
0	1	2	3	4	5
1	Sausis . . .	4 146	3 088	3 206	3 480
2	Vasaris . . .	3 598	2 829	2 934	3 120
3	Kovas . . .	3 472	3 078	3 400	3 317
4	Balandis. . .	2 954	2 876	2 949	2 926
5	Gegužis . . .	2 886	2 656	2 878	2 807
6	Birželis . . .	2 508	2 625	2 255	2 463
7	Liepa. . . .	2 254	2 963	2 212	2 477
8	Rugpiūtis . .	2 281	3 712	2 405	2 799
9	Rugsėjis. . .	2 169	2 986	2 256	2 470
10	Spalis . . .	2 366	2 863	2 322	2 517
11	Lapkritis . .	2 744	2 705	2 518	2 656
12	Gruodis. . .	3 217	3 408	3 414	3 346
		34 595	35 789	32 749	34 378

Tą patį reiškinį mes matome ir mirimų statistikoj, — nors patys mėnesiniai svyravimai čia mažesni. Matyti, žmonės nė jų institucijos dar neišgalvojo priemonių, kurios įgalintų atskirais mėnesiais savivalingai mirimų skaičių sumažinti. Žymes-

nių skirtumų čia galima pastebėti tik pirmiesiems žiemos mėnesiams, kai mirtingumas būna didžiausias: tik vienais metais mirimų skaičius tenka gruodžio mėn., o kitais — sausio mėn.

Arba imkime kitą pavyzdį iš Lietuvos gimimų statistikos. Jau S u e s s m i l c h'as stebėjosi nepaprastu dėsningumu, kuris pasireiškia gimstančiųjų berniukų ir mergaičių santykyje. 20-čiai mergaičių, aplamai, esą gimsta 21 berniukas. Bet šis dėsnis pasireiškiąs tik gimimų masėje. Atskirose šeimose, arba net ir atskirose vietovėse, esą kai kada gimsta vieni berniukai, tai vėl vienos mergaitės — arba jau labai nelygioj proporcijoj.

Šį dėsningumą mes galime matyti ir L i e t u v o s gimimų statistikoj:

Lentelė 30.

GIMIMŲ SKAIČIAI LIETUVOJE.

	Berniukai	Mergaitės	100-tui gim. mergaičių gimė berniukų:
1922	30 149	27 915	108
1923	31 289	29 580	106
1924	32 802	31 062	106
1925	32 574	31 169	105
1926	32 724	30 931	106
1927	34 118	31 996	106
1928	33 847	32 098	105
1929	32 643	30 440	107
1930	32 976	31 188	106
1931	32 691	30 728	106
1932	33 660	31 711	106
1933	31 999	30 146	106
1934	31 247	29 523	106
1935	29 821	28 149	106
1936	31 099	29 347	106

Kaip matome, per 15 metų 11 kartų pasikartoja skaičius 106, du kartu 105, vieną kartą 107 ir vieną kartą 108. Bet jeigu paimsime tuos pačius duomenis, paskirstytus apskritimis, tai jau rasime gana didelių nukrypimų. Pav., 1931 m., kai bendrai Lietuvoje 100-ui mergaičių buvo gimę 106 berniukai, Panevėžio

apskrityj terandame 99 berniukus, Zarasų apskr. tik 96, užtat Panevėžio mieste 118, Ukmergės mieste net 123. O gal būt tose vietose visuomet tokia proporcija būva? Pasirodo, kad ne visuomet. Pav., Panevėžio apskr. 1930 m. 100-ui mergaičių gimė 103 berniukai, 1929 m. net 106 ir t. t. Šiaulių apskr., kur 1931 m. 100-ui mergaičių buvo gimę tik 99,7 berniukai, 1930 m. berniukų koeficientas pakyla net ligi 114, 1929 m. vėl nukrinta ligi 99. Klaipėdos mieste vienais metais 100-ui mergaičių buvo gimę net 119 berniukų, ir t. t.

Šie pavyzdžiai rodo, kad vienos ar kitos rūšies skaičių pastovumas, arba vad. „statistinis dėsningumas", yra didesnis, kai stebimoji masė gausingesnė, ir atvirkščiai: nuo bendrojo dėsnio (nuo normos) didesni nukrypimai dažniau pasireiškia mažesnėje masėje. Žodžiu, empiriškas gyvenimo reiškinių stebėjimas visai patvirtina vad. didžiųjų skaičių dėsnį.

Tačiau šių laikų žmogui dar nepakanka žinoti, kad viena ar kita teorija gyvenimo patirimams neprieštarauja, arba kad tie patirimai teorijos apibrėžtąjį dėsnį patvirtina. Maža žinoti, kad vedybų, mirimų ar kurio nors kito reiškinio tam tikros ypatybės, arba skirtumai, pasireiškia tik metiniuose arba mėnesiniuose duomenyse — arba imant tik visą valstybės plotą, o ne atskiras jos dalis. Nedaug dar mums padės ir tai, jei sužinosime, kad toki skirtumai gali pareiti nuo įvairių priežasčių, iš kurių vienos gali būti atsitiktinės (pav. vienais metais pavasaris vėliau, kitais — anksčiau), kitos struktūrinės (tam tikros masės dalinio stebėjimas, kurios nors ypatingos sąlygos).

Statistikui šie du skirtingi vieno ir to paties darbo (metinių duomenų suskaidymo į mėnesinius arba didesnio ploto į smulkesnius) rezultatai diktuote diktuoja kitą klausimą, būtent: jei vienodu dėsningumu pasikartojančius duomenis pertvarkius didesniais ar mažesniais kiekiais, nuo „normos" (vidurkio) nukrypimai pasirodo didesni, arba jeigu atsitinka, kad su buvusiais dėsningumais panašumo nelieka, — tai ar nėra metodų bei priemonių šiek tiek griežčiau ir tiksliau nustatyti, kuriais atvejais masinio stebė-

jimo duomenyse didžiųjų skaičių dėsnio veikimas pasireiškia, o kuriais atvejais yra, arba gali būti, pagrindas tarti, kad didžiųjų skaičių dėsnio veikimas dar nepasireiškė? Kitais žodžiais: kaip surasti ir nustatyti kiekvienu atskiru atveju, kada statistiniam stebėjimui bei vaizdavimui paimtųjų vienetų skaičius yra pakankamas, kad didžiųjų skaičių dėsnis pasireikšti galėtų, o kada tasai skaičius dar nepakankamas? Žodžiu, iš ko pažinti, kad empirinio stebėjimo keliu gautųjų duomenų rodomi vienos ar kitos rūšies tariamieji dėsningumai nėra atsitiktiniai, nieko bendra neturintieji su didžiųjų skaičių dėsningumais?

b) Atsitiktinių įvykių eigos dėsningumai.

Šioj srityj statistikai labai daug patarnavo prof. Westergaard'as savo tyrinėjimais. Be visos eilės kitų tyrinėjimų, prof. Westergaard'as paėmė po lygų skaičių visai vienodo didumo ir sunkumo raudonai ir baltai dažytų rutuliukų ir, supylęs į dėžutę bei gerai juos išmaišęs, vieną atsitiktinai į rankas patekusį rutuliuką išėmė ir čia pat jo spalvą pasižymėjo. Paskui jis tą rutuliuką ir vėl atgal į dėžutę įdėjo ir, prieš sekantį traukimą, vėl dėžutėje esančius rutuliukus išmaišė, o po to rutuliuko traukimą pakartojo. Ir t. t. Tuo būdu rutuliukų traukimas buvo padarytas visuomet tose pačiose sąlygose, t. y. esant dėžutėje po lygų abiejų spalvų rutuliukų skaičių. Kurios spalvos rutuliukas kada ištraukiamas — nulemia aklas likimas. Teoriškai galima šitaip galvoti: kadangi, traukiant iš krepšelio ar dėžutės rutuliuką, kiekvienu atveju būna visai vienodas ir lygus raudonų ir baltų rutuliukų skaičius, tai, rodos, galėtų atsitikti, kad bus pataikyta ištraukti vis kurios nors vienos spalvos rutuliukas — arba bent jau vienos spalvos rutuliukas gali iš eilės pavykti ištraukti keliasdešimtį kartų. Tuo būdu, bent teoriškai, nėra pašalinta galimybė, sakysim, 100 arba ir 1000 kartų iš eilės ištraukti vis baltus ir baltus (arba tik raudonus) rutuliukus. Tačiau, iš kitos pusės, yra teoriška galimybė pataikyti kas antrą raudoną rutuliuką, — kadangi abiejų spalvų rutuliukų skaičių santykis yra visai vienodas, lygus

1-nui. Ir tuo būdu, grynai teoriškai galvojant, rutuliukų traukimą žymimos raidžių eilutės gali atrodyti:

bbbbbbbbbbbbbb....., arba
rrrrrrrrrrrrrr....., arba
brbrbrbrbrbrbrbr....., arba kaip nors kitaip, pav.,

bbrbbbrrrrbrbbrrrbb....., ir t. t.

Prof. Westergaard rutuliukų traukimą pakartojo 10 000 kartų, ir 5 011 kartų pataikė ištraukti baltą rutuliuką, o 4 989 kartus raudoną. Taigi visai vienodose sąlygose imant bet kurį atsitiktinai į ranką patekusį rutuliuką, ištrauktųjų rutuliukų skaičių santykis maža kuo tesiskyrė nuo faktinai buvusio rutuliukų skaičių santykio. Nes 5 011 : 4 989 (prie buvusios 5 000 : 5 000) išeina lygybės, jei pasitenkinti mažesniu griežtumu, pav., kad ir sveikais procentų skaičiais. Ir tik siekiant didesnio griežtumo, pav., norint eksperimento rezultatus pareikšti tūkstantinės dalies griežtumu (promiliais) gaunama nelygybė. Westergaard'o darytųjų bandymų duomenis suskirsčius šimtinėmis pagal raudonųjų ir baltųjų rutuliukų skaičių, kiekvienoj šimtinėj buvo gauta (žiūr. psl. *242*):

Jei imti šiuos prof. Westergaard'o stebėjimų duomenis kaip bet kurios statistinės masės reprezentantus, tai statistinio analizio uždavinys gali būti dvejopas: 1) ištirti bei apibrėžti dispersijos pobūdį, ir *2*) nustatyti, ar atskiros reprezentantų grupės bei jų sudėtis atitinka normas, kurias aplamai apibrėžia atsitiktinio stebėjimo arba atsitiktinių įvykių tikimybės teorija.

Pirmajam uždaviniui išnagrinėti mes galime imti tuos pačius duomenis ir sulig tuo, kiek kurioj šimtinėj buvo vienos spalvos rutuliukų, sudaryti x-inės eilutės dydžių grupes, kurių pasikartojimo skaičiai atstovaus y-inės eilutės skaičius. Vidurinis kvadratingasis nukrypimas, taip pat dispersijos koeficientas, turi parodyti, normali ar nenormali yra stebėjimo duomenų dispersija. Šis gi nukrypimas, pagal bet kurią formulą išskaičiuotas griežtumu ligi 0,1 duoda $\sigma\,(=)\,5{,}2$. Taigi dispersijos koeficientas, kai „norma" (vidurkis) sudaro apie 50 (griežtai baltiems rutuliukams skaičiuojant 50,11), išeina $d_\sigma = \frac{\sigma}{\tilde{x}} =$

Lentelė 30a.

WESTERGAARD'O TYRINĖTŲJŲ DVIEJŲ VIENODAI GALIMŲ ĮVYKIŲ ATSITIKTINIAI NUO „NORMOS" NUKRYPIMAI, KAI „NORMOS" b IR r SANTYKIAUJA KAIP 50 : 50.

Iš 100 grupių buvo gauta:

Eil. nr.	Grupių skaičius.	Rutuliukų skaičius kiekvienoj šimtinėj. baltų	raudonų
(n)	(y)	(b)	(r)
0	1	2	3
1	1	34	66
2	1	39	61
3	2	40	60
4	2	41	59
5	2	42	58
6	3	43	57
7	3	44	56
8	4	45	55
9	5	46	54
10	6	47	53
11	5	48	52
12	11	49	51
13	9	50	50
14	5	51	49
15	10	52	48
16	4	53	47
17	8	54	46
18	3	55	45
19	5	56	44
20	4	57	43
21	4	58	42
22	1	61	39
23	1	62	38
24	1	63	37

$= \frac{5,2}{50,11}$ (=) 0,1 arba 10%. O mes jau žinome, kad jeigu dispersijos koeficientas neprašoka 50%, tai st. eilutė dar gali būti laikoma homogenine, o aritmetinį vidurkį galima laikyti atstovaujančiu visą st. eilutę.

Medianas šioj st. eilutėj M = 50. Kai a. vidurkis $\bar{x}$ (=) 50,1 ir $\sigma = 5,2$, tai asimetrijos koeficientas $a_\sigma = \frac{3(\bar{x} - M)}{\sigma} =$ $= \frac{3(50,1 - 50,0)}{5,2} = \frac{3.0,1}{5,2} = \frac{0,3}{5,2}$ (=) 0,06. Tuo būdu ir asimetrijos koeficientas rodo, kad dispersija gali būti artima normaliai.

Normalios dispersijos griežtesnį apibrėžimą duoda vad. t r i j ų s i g m ų taisyklė (žiūr. psl. 212). Kai $\sigma = 5,2$, o $\bar{x} = 50,1$, tai $\bar{x} \pm 3\sigma$ ribose esant normaliai dispersijai, sutelpa apie 99,7% visų st. eilutės vienetų, $\bar{x} \pm 2\sigma$ — apie 95,5% ir $\bar{x} \pm 1\sigma$ — apie 68,3%. Mūsų nagrinėjamame atsitikime trijų sigmų intervalo rėžiai nustatomi toki:

1. $50,1 \pm 1 . 5,2 = 44,9 — 55,3$
2. $50,1 \pm 2 . 5,2 = 39,7 — 60,5$
3. $50,1 \pm 3 . 5,2 = 34,5 — 65,7$

Iš 31 lentelės matome, kad už rėžio 65,7 vienoj šimtinėj, nėra nė vieno atsitikimo, o už 34,5 — tik 1 iš 100 grupių. Vadinasi, iš visų grupių 99% sueina trijų sigmų intervalo rėžiuose. O turint galvoje, kad iš viso baltųjų rutuliukų ištraukta 5 011, ir jų tarpe 1 šimtinėj pasireiškęs didžiausias nuo normos (50,1) nukrypimas sudarė 34, tai, skaitant procentais, šis nukrypimas tesudaro tik $\frac{34}{5011}$ (=) 0,0067 arba apie 0,67%. Vadinasi, iš tikrųjų trijų sigmų intervalo rėžiuose sueina 100,00—0,67 = 99,33%, t. y. beveik tiek pat, kaip prie griežtai normalios dispersijos.

Dviejų sigmų riba sudaro intervalą nuo 39,7 ligi 60,5. Už šio intervalo ribos iš viso tėra penkios šimtinės (su pažymiais 34, 39, 61, 62 ir 63), kur b rutuliukų bendroji suma siekia 259, taigi sudaro $\frac{259}{5\,011}$ (=) 0,052 arba 5,2%. Tuo būdu dviejų kvadratingųjų nukrypimų ribose palieka 100,0 — 5,2 = 94,8%.

Vieno kvadratingojo nukrypimo riba aprėžia intervalą nuo 44,9 ligi 55,3 b_i. Už šio intervalo ribų yra palikę iš viso 30 šimtinių su baltųjų rutuliukų bendrąja suma, siekiančia 1 506 v. iš 5 011. Ši suma sudaro taip pat apie $30\%\left(\frac{1\,506}{5\,011}\ (=)\ 0{,}3006\right)$. Taigi vienos sigmos ribose telpa net $100 - 30 = 70\%$, — kai griežtai normali dispersija privalo arti 68,3%.

Visi šie duomenys duoda pagrindą teigti, kad empiriškai stebėtieji atsitiktiniai įvykiai yra pasiskaidę beveik tuo pačiu dėsningumu bei proporcingumu, kaip ir matematiškai galimųjų tokio pat skaičiaus įvykių kombinacijos, žinomos iš binomo koeficientų pasiskaidymo eilutės. Lieka tada tik išnagrinėti teoriškojo ir empiriškojo pasiskaidymo galimų skirtumų tikimybes.

c) Atsitiktinių įvykių matematiškoji tikimybė.

Štai eina sau ramiausiai, kaip ir kasdien, Jonas skersai Laisvės alėją ir patenka po automobiliu. Dėl ko būtent ši nelaimė palietė Joną, o ne Petrą, kuris taip pat tuo pačiu keliu ir tuo pačiu laiku vaikščioja? Dėl ko tą patį Joną ištiko nelaimė šiandien, o ne vakar, ne prieš metus laiko? Ir t. t. Arba štai du draugai vienu kartu perka loterijos bilietą. Bet vieno bilietas „išlošia", o antrojo — „tuščias" palieka. Kas čia? — Vėl atsitiktinumas — „lemtis", „likimas" — atsakoma. Bet vis dėlto yra žinoma, kad jeigu iš 100 loterijos bilietų 50 yra „tušti", tai tas, kuris paims 60 bilietų, būtinai turės gauti bent 10 „pilnų" — kad ir visa 50 tuščiųjų pataikytų. Arba tarkime, kad mes žinome, jog Kauno mieste yra 20 000 gyventojų, kurie iš viso turi per metus 40 mil. litų darbo pajamų, — vadinasi, vidutiniškai po 2 000 litų kiekvienas. Klausimas — kiek jų tarpe yra tokių kauniečių, kurių metinės pajamos nėra mažesnės kaip 10 000 litų. — Aišku, kad tokių žmonių gali būti m a ž i a u kaip 4 000: nes jeigu 4 000 piliečių turėtų po 10 000 lt. pajamų, tai likusiems 16 000 neliktų nieko — ir jie būtų nepatekę į tą 20 000 skaičių.

Vadinasi, ir „likimo" ribos tam tikromis priemonėmis duodasi apibrėžiamos. Taigi ir „nežinomųjų" skaičius ligi tam tikro laipsnio duodasi aprėžiamas. O jeigu taip, tai ir patsai at-

sitiktinumas, arba atsitiktinių įvykių eiga, gali būti stebima ir vad. „likimo“ arba „lemties“ dėsningumų ieškoma.

Aplamai, bet kurios rūšies dydžių eilutės pasiskaidymą stebint, tenka pasistatyti ir ieškoti atsakymo į šiuos klausimus: 1) kokio pobūdžio ir intensyvumo yra dispersija, t. y. dydžių pasiskaidymas apie vidurkį, 2) kuris iš visų dydžių gali pasikartoti dažniausia, t. y. kurio, būtent, įvykio arba įvykių kombinacijos, tikimybė didžiausia ir 3) kaip dideli gali būti nukrypimai nuo „normos“ arba „laukiamos“ tikimybės, t. y. kaip didelė gali būti patikimoji (taip pat vad. normali ir galimoji) paklaida.

Šie trejopo pobūdžio klausimai gali būti atsakyti tam tikromis matematinio analizio priemonėmis. Pav., į pirmosios kategorijos klausimus duoda pakankamai aiškų atsakymą kvadratingasis nukrypimas bei dispersijos ir asimetrijos koeficientai. Bet antrosios ir trečiosios kategorijų klausimai, kad ir susieti su pirmuoju, reikalingi papildomų duomenų.

Visų pirma, k u r i o, būtent, į v y k i o, arba įvykių kombinacijų, duotomis sąlygomis t i k i m y b ė v i s ų d i d ž i a u s i a? — Šitaip pastatytam klausimui pilnumoj atsakyti tegali tik tikimybių teorija. Statistikui, tačiau, pakanka įsisąmoninti tai, kas jau žinoma iš binominio pasiskaidymo dėsningumų. Mes jau matėme, kad vieną kartą traukiant rutuliuką iš dėžutės, kur yra baltų ir raudonų rutuliukų po lygų skaičių, ištraukti raudoną rutuliuką tikimybė $p = \frac{1}{2}$ ir neištraukti — $q = \frac{1}{2}$. Vadinasi, šiuo atveju „atspėjimo“ ir paklaidos tikimybės yra vienodos. Bet kas atsitiks, jeigu bus traukiama ne vieną kartą, o, sakysime, 2, 5, 10, 100, 10 000 ir dar daugiau kartų? Ar, sakysim, traukimą kartojant 100 ar 100 000 kartų, tikimybės — pataikyti kiekvieną kartą baltą arba raudoną rutuliuką — paliks visai vienodos?

Vėl grįžkime prie binominio pasiskaidymo normų bei dėsningumų. Visų pirma, mūsų pažįstamoji binominės lygties kairioji pusė, kuri paprastai rašoma $(a + b)^n$, nė kiek savo turinio nepakeis, jei narį a pakeisime laukiamos ar pageidaujamos tikimybės simboliu p, o antrojo nario simbolį b parašysim ne-

laukiamos tikimybės (paklaidos) simboliu q. Vadinasi, binomo formulą mes galime ir šitaip parašyti:

$$(a + b)^n = (p + q)^n.$$

Toliau mes žinome, kad jeigu a ir b yra natūriniai bei vienareikšmiai skaičiai, kurie santykiauja kaip 1 : 1, tai mūsų formula duosis ir šitaip parašoma:

$$(p + q)^n = (a + b)^n = (1 + 1)^n = 2^n.$$

Jei sutarsime apibrėžtomis sąlygomis d u kartu mėginti „laimės“ (ištraukti rutuliuką, mesti pinigą ir laukti atsiverčiant skaitmens), tai g a l i m ų įvykių kombinacijų, išreikštų jų tikimybes išeis:

aa, ab, ba, bb (arba: pp, pq, qp, qq).

Bet kai $n = 2$, $a = b$, $a = 1$ ir $a + b = 2$, tai mūsų formuloj:

$$(a + b)^n = (1 + 1)^2 = 4.$$

O jeigu sutarsime p, q žymėti natūrinių *a, b* skaičių tikimybes, tai $p + q = 1$ ir $p = \frac{1}{2}$, nes: kai $a + b = 2$ ir $a = b = 1$, tai

$$pa = \frac{a}{a + b} = \frac{1}{1 + 1} = \frac{1}{2}.$$

Tuo pačiu būdu gaunama tikimybė q. Tuo būdu galime bet kurio laipsnio tikimybes p, q žymėti p^n, q^n. Kai bandymų darome dvejetą, t. y. $n = 2$, tai $p^n = p^2$, vadinasi: $p^2 = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$. Šiame bandyme tie patys rezultatai gausis ir priešingam atvejui q, vadinasi, ir $q^2 = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$. Bet bandyme galime gauti ir pq arba qp (panašiai kaip ab ir ba). O kadangi pq ir qp rezultate yra tapatybės, tai užuot rašyti: kad

$$(p + p)^2 = p^2 + q^2 + pq + qp$$

galime rašyti:

$$(p + q)^2 = p^2 + q^2 + 2\,(pq).$$

O kai sutarėme, kad $(p + q)^2 = 1$, tai, simbolius pakeitę skaičiais, gausime:

$$(p + q)^2 = \left(\frac{1}{2} + \frac{1}{2}\right)^2 = 1.$$

Jei bandymą pakartosime tris kartus, vadinasi, n = 3, tai

$$(p + q)^3 = p^3 + q^3 + {}^3pq\,(p + q) = \frac{1}{8} + \frac{1}{8} + 3\,\frac{1}{2}\,\frac{1}{2}\left(\frac{1}{2} + \frac{1}{2}\right) =$$

$$= \frac{1}{8} + \frac{1}{8} + \frac{3}{4}\left(\frac{1}{2} + \frac{1}{2}\right) = 1.$$

Ir t. t.

Žodžiu, bandymų arba potyrių (stebėjimų) skaičiui (n) didėjant, galimų ir negalimų įvykių tikimybė (tikrenybės laipsnis), tiek jų eilutės sekoj galimų kombinacijų skaičius vyksta binominės eilutės narių išdėstymo dėsningumu (žiūr. psl. 210, arba atitinkamą skyrių iš algebros vadovėlio). Jei kas pravartu įsidėmėti, tai tik aplinkybė, kad binomo išvystymo ir jo koeficientų pasiskaidymo matematiškomis priemonėmis apibrėžimas priklauso begaliniam skaičiavimui. Statistikoj gi, kuri turi reikalą tik su apibrėžtomis arba bent apibrėžiamomis masėmis, be galo didelių nei be galo mažų skaičių, kaip tokių, nėra. Ši aplinkybė įgalina statistines tikimybes paprastesnėmis matematinio analizio priemonėmis ir daugiau artutiniais dydžiais pareikšti, ir, be to, normalios ir net galimos paklaidos ribas griežčiau apibrėžti. Pav., binomo koeficientų dispersija, apibrėžiama kvadratinguoju nukrypimu (σ), jau prie $\pm 3\sigma$ duoda intervalo ribas, kuriose sueina 99,73% visų koeficientų; tačiau nei $\pm \sigma 3\,000$, nei bet kuris dar didesnis sigmų skaičius neduos intervalo ribų, kur sutilptų visas 100%. Tuo tarpu stebint statistines mases, jei dispersija bent artutinė normaliai, jau $\pm 3\sigma$ ribose dažniausia pasiekiama 100%, t. y. tikrenybės, ir tik retais atsitikimais patys kraštutiniai nukrypimai prašoka $\pm 4\sigma$ ribą.

Šis skirtumas ypač lengva suprasti turint prieš akis Westergaard'o tyrinėjimus. Traukiant rutuliukus iš dėžės, kur visuomet yra po lygų skaičių abiejų spalvų rutuliukų, sudaromos statistinio stebėjimo sąlygos begalybėje (faktiškai niekad negalima traukimo, stebėjimo „baigti", jei ištrauktasis rutuliukas kiekvieną kartą vėl atgal įdedamas). Ir Westergaard'as, šimtinėmis stebėjimo rezultatus suskaidęs, kad ir gavo keletoj

šimtinių faktinąjį baltų ir raudonų rutuliukų skaičių santykį, bet tai nieko nereiškia: nes, visų pirma, patsai šimtinių sudarymas yra savivalingas darbas (kodėl šimtinėmis, o ne penkšimtinėmis arba ne dešimtinėmis grupėmis?), o antra — ir 10 000 kartų traukimą pakartojęs, baltuosius ištraukė 5 011 kartų, o raudonuosius tik 4 989. Vadinasi, paklaidą sudarė ± 0,0011, t. y. ± 0,22%. O jeigu dėžėj būtų buvę tik 5 000 baltų ir 5 000 raudonų rutuliukų, bet ištrauktųjų rutuliukų atgal nebūtų įdėta, tai 10 000 kartų ištraukus būtų pasiekta absoliuti tikrenybė be jokios paklaidos. Vadinasi, statistikos praktikoj, kai n yra baigtinis ir pakankamai didelis skaičius, patikimoji ir ypač galimoji paklaida, apibrėžiama begalinio analizio priemonėmis, sudaro kraštutines nuo normos nukrypimo ribas, t. y. praktiškai parodo, kur baigiasi tikimybė ir prasideda tikrenybė. Jei urnoje yra 5 000 baltų ir 5 000 raudonų rutuliukų, tai ištrauktųjų raudonų rutuliukų skaičius negali būti didesnis kaip 5 000 net ir tuo atveju, jei traukiama tik 5 000 kartų, o ne 8 000 ar 10 000 (jei n = 10 000).

Iš kitos pusės, labai maža yra vilties ištraukti iš urnos visus 5 000 raudonųjų rutuliukų pirmaisiais 5 000 traukimų, t. y. labai maža vilties, kad iš urnos, kur buvo supilta 5 000 raudonų ir 5 000 baltų rutuliukų, iš eilės 5 000 kartų a t s i t i k t i n a i būtų pataikyta vis raudoną ir raudoną rutuliuką ligi pat paskutiniojo. Aplamai, matematiškai ir empiriškai (to paties Westergaard'o ir kt.) yra nustatyta, kad maksimalinė tikimybė mažėja, potyrių (bandymų) skaičiui n didėjant, ir net yra nustatyta maksimalinės tikimybės mažėjimo dėsnis. Mažėjimas vyksta pagal proporciją $\frac{1}{\sqrt{n}}$ (vadinamas kvadratingosios šaknies dėsnis).

Matematikai turi formulų, kurios mažesniu ar didesniu griežtumu apibrėžia maksimalinę tikimybę. Mums rūpimoj normalioj dispersijoj (binomo išsivystyme) maksimalinės tikimybės skaičiui statistikoj (kai n yra pakankamai didelis skaičius) galima panaudoti sutrumpintą Stirling'o formulą, kur

$$P_0 = \frac{1}{\sqrt{2\pi npq}} \qquad \text{(XXVIII).}$$

Kai π yra iš geometrijos žinomoji konstanta 3,14159..., kurią statistikos reikalui drąsiai galima trumpinti ligi šimtinės dalies griežtumo ir rašyti $\pi\,(=)\,3{,}14$, tai binominiams dydžiams koeficientų maksimalinė tikimybė P_0, jeigu $n = 10$, sudarys apie $\frac{1}{4}$.

Nes kai $p = \frac{1}{2}$, o $p + q = 1$, ir $n = 10$, tai pagal formulą XXVIII

$$P_0 = \frac{1}{\sqrt{2\pi npq}} = \frac{1}{\sqrt{2 \,.\, 3{,}14 \,.\, 10 \,.\, 0{,}5 \,.\, 0{,}5}} =$$

$$= \frac{1}{\sqrt{6{,}28 \,.\, 2{,}5}} = \frac{1}{15{,}45} \;(=)\; \frac{1}{3{,}93} \;(=)\; 0{,}25.$$

Iš psl. 215 lentelės 19 galime įsitikinti, kad iš tikrųjų binomo koeficientas $(n)_k = (10)_5$ formuloj $(a + b)^n = 2^{10}$ raidoje gaunamas 252, kai bendroji suma $2^{10} = 1\,025$.

Nes:

$$\binom{n}{k} = (n)_k = \frac{n!}{k!\,(n-k)!},$$

kai $\binom{n}{k} = \binom{10}{5}$, išeina:

$$\binom{10}{5} = \frac{10!}{5!\,(10-5)!} = \frac{1 \,.\, 2 \,.\, 3 \,.\, 4 \,.\, 5 \,.\, 6 \,.\, 7 \,.\, 8 \,.\, 9 \,.\, 10}{1 \,.\, 2 \,.\, 3 \,.\, 4 \,.\, 5 \,.\, 5 \,.\, 4 \,.\, 3 \,.\, 2 \,.\, 1} =$$

$$= \frac{3\,628\,800}{14\,400} = 252.$$

O

$$252 : 2^{10} = \frac{252}{1025} \;(=)\; 0{,}25$$

Čia vad. Stirling'o formulos veikimas, patikrintas iš algebros žinomo binomo koeficientų formulos praktišku pritaikymu, rodo tokį didelį sutapimą, kad yra pagrindo teigti, jog ir kitais atsitiktinių įvykių atvejais (masinius reiškinius stebint bei tyrinėjant) galima apibendrintomis matematikos formulomis operuoti bei jomis pasitikėti.

Visų pirma, pravartu įsidėmėti, kad tai, ką matematikai vadina dydžiu su maksimaline tikimybe, pažymima

P_0, W_0, E ar dar kaip nors kitaip, ir kur lygties priešingoj pusėj kai kada prirašo įvairiausiais pažymiais apibrėžtų raidžių, skaičių ar kt. ženklų, statistikoj vadinama dažniausia pasitaikančiu dydžiu arba dominantu. Ir jeigu st. eilutės dydžių dispersija yra bent artutinė normaliai dispersijai, tai, kaip žinoma, dominantas perdaug neatsitolina nei nuo mediano, nei nuo aritmetinio vidurkio. Vadinasi, čia patiektoji sutrumpinta Stirling'o formula yra tik pagalbinė priemonė griežčiau apibrėžti vidurkinio dydžio pasikartojimą arba jo lyginamąjį svorį visoje masėje tais atvejais, kai st. masė yra labai didelė ir jos kokybinė klasifikacija smulkesniais vieno dydžio intervalais neįmanoma, — arba kai st. masės stebėjimas daromas reprezentaciniu metodu.

Bet šioj vietoj pravartu susipažinti ir su matematikos teikiamomis priemonėmis tikimybinių dydžių nukrypimams nuo „normos", t. y. paklaidoms, nustatyti. Nes statistikos praktikoj dar maža žinoti, kad dominantas, aritmetinis vidurkis ar kuris nors kitas st. eilutės x_i dydis gali sudaryti apie tiek procentų arba promilių visoje masėje. Kai kada gali būti svarbiau žinoti — kiek gi vidurkis ar kuris nors kitas dydis sau a r t i m e s n i ų dydžių gali turėti. Vadinasi, reikalinga dar statistinio analizio priemonė, kuri įgalintų bet kurio st. masės dalinio, ar jį reprezentuojančio dydžio, vad. normalią paklaidą apibrėžti panašiai, kaip viduriniu kvadratinguoju nukrypimu apibrėžiama aritmetinio vidurkio santykiavimas su visa mase.

Iš tikrųjų tokia priemonė jau yra numatyta Stirling'o maksimalinės tikimybės formuloje XXVIII. Jei paėmę tą formulą atpalaiduosime ją nuo reiškinio 2π, tai gausime bet kurios tikimybės p narių skaičiaus kvadratingąjį nukrypimą, kuris, paprastai, taip pat žymimas mažąja sigma ir gaunamas pagal formulą:

$$\sigma = \sqrt{npq} \qquad \text{(XXIX)}.$$

Šios sigmos reikšmė tikimybiniam dydžiui, event. nuo „normos" nukrypimams, yra ta pati kaip ir sigmos, gautosios pagal bet kurį XXVII formulos variantą.

Vėl visų pirma patikrinkime šios formulos veikimą binominėj eilutėj.

Jei a = b, a + b = 2, ir n = 1, tai a gauti tikimybė $p_a = \frac{1}{2}$, o tuo būdu ir $p_b = 1 - \frac{1}{2} = \frac{1}{2}$. Šiomis sąlygomis vidurinis kvadratingasis nukrypimas taip pat bus $\sigma = \frac{1}{2}$, nes:

$$\sigma = \sqrt{npq} = \sqrt{1 \,.\, 0{,}5 \,.\, 0{,}5} = 0{,}5.$$

Tačiau, jeigu n = 10, tai, p ir q dydžių santykiui nepasikeitus, ta pati formula duos apie 1,57, nes:

$$\sigma = \sqrt{10 \,.\, 0{,}5 \,.\, 0{,}5} = \sqrt{2{,}5}\ (=)\ 1{,}57.$$

Kaip matome, šiuo atveju gautasis σ visai atitinka σ, gautą iš binomo koeficientų eilutės psl. 249, kur $(a + b)^n = 2^{10}$, taigi potyrių skaičius n = 10. O ten buvo panaudota formula XXVII, kur $\sigma = \sqrt{\frac{\triangle^2}{n}}$.

Taigi galime konstatuoti, jog:

$$\sigma = \sqrt{npq} = \sqrt{\frac{\Sigma \triangle^2}{n}}.$$

Vadinasi, vidurinis kvadratingasis nukrypimas yra patikimas rodiklis ne tik st. eilutės dydžių dispersijai ir vidurkiui charakterizuoti, bet ir bet kurio st. eilutės dydžio x_i ir jį atitinkančio dalinio gaunamam koeficientui griežčiau apibrėžti. — Nes tikimybinis dydis p, paimtas iš statistinės eilutės, atstovauja ne ką kitą, kaip mums jau iš 137 psl. pažįstamą statistinio santykio dalinį d_i, kuris taip pat gali apimti 0,1, 0,5, 0,7 ar kurią nors kitą visumos dalį. Ir jeigu tam tikromis sąlygomis turėdami prieš akis $\sigma = \sqrt{\frac{\triangle^2}{n}}$ galime teigti, kad apie aritmetinį vidurkį vieno kvadratingojo nukrypimo ribose yra nemažiau kaip ⅔ visų st. eilutės dydžių, tai tuo pačiu pagrindu galime sakyti, kad $\sigma = \sqrt{npq}$ ribose, prie normaliai artutinės dispersijos, taip pat apie ⅔ visų dalinių grupuojasi apie tikimybinį p dydį. Vadinasi, bet kuris st. eilutės dydis x_i, sudarydamas tikimybės p apibrėžtą dalinį visame vienete, yra tokiame pat

santykyje su $\sigma = \sqrt{npq}$, kaip x-inės eilutės vidurkis $\bar{x}$ su $\sigma = \sqrt{\frac{(x_i - \bar{x})^2}{n}} = \sqrt{\frac{\Sigma \triangle^2}{n}}$. Iš to seka, kad tikimybės (p) kvadratingasis nukrypimas gali būti gautas ir iš modifikuotos formulos, pav.,

$$\sigma = \sqrt{\frac{pq}{n}}. \qquad \text{(XXIXa).}$$

O jeigu p = q, kaip yra binominėj eilutėj, tai galima rašyti ir

$$\sigma = \sqrt{\frac{p^2}{n}}.$$

Nes jeigu p = 0,5 ir q = 1 — p = 0,5, o n = 1, tai visais atvejais $\sigma = 0{,}5$:

$$\sigma = \sqrt{npq} = \sqrt{1 \,.\, 0{,}5 \,.\, 0{,}5} = \sqrt{0{,}25} = 0{,}5;$$

$$\sigma = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0{,}5 \,.\, 0{,}5}{1}} = \sqrt{0{,}25} = 0{,}5.$$

$$\sigma = \sqrt{\frac{p^2}{n}} = \sqrt{\frac{0{,}5^2}{1}} = \sqrt{0{,}25} = 0{,}5;$$

Tie patys rezultatai bus ir prie n = 10, nes:

$$\sigma = \sqrt{\frac{0{,}5 \,.\, 0{,}5}{10}} = \sqrt{\frac{0{,}25}{10}} = \sqrt{0{,}025}\ (=)\ 0{,}157.$$

Šis σ skaičius visai atitinka psl. 251 gautąjį, nors tenai buvo gauta $\sigma\ (=)\ 1{,}57$, o čia $\sigma\ (=)\ 0{,}157$, t. y. trupmena. Norėdami gauti kvadratingąjį nukrypimą σ taip pat sveikais racionaliais skaičiais, turime ir tikimybinius dydžius p pareikšti ne santykio daliniais, o absoliučiais skaičiais. Šiuo atveju, kai n = 10, o tikimybės p ir q yra lygios, tai tikimybinius dydžius p, q atstovaują absoliutūs skaičiai a, b bus: a = 10 — b = 5 ir b = 10 — a = 5. Vadinasi, galime rašyti:

$$\sigma = \sqrt{\frac{5 \,.\, 5}{10}} = \sqrt{\frac{25}{10}} = \sqrt{2.5}\ (=)\ 1{,}57.$$

Iš čia matome, kad, galų gale, griežtai normalios dispersijos eilutėj ne tik

$\sigma = \sqrt{\frac{\Sigma \triangle^2}{n}} = \sqrt{npq}$, bet ir $\sqrt{\frac{pq}{n}}$ duoda tą patį skaičių, jei tikimybinius dydžius p, q pakeisti jų originaliniais skaičiais a, b.

Kuris kvadratingojo nukrypimo formulos variantas naudotinas — pareina nuo aplinkybių. Aplamai, formulos XXIXa variantas naudotinas visais tais atvejais, kai norima iš karto gauti dispersijos koeficientas, t. y. vidurinio kvadratingojo santykinis skaičius. Nes kai p ir q atstovauja tikimybių skaičius (trupmenas), kurių suma lygi 1, tai σ iš formulės $\sqrt{\frac{pq}{n}}$ taip pat tegali duoti tik trupmenas.

Vėl paimkime tuos pačius binominės eilutės dydžius. Jei binomo lygtyj n = 100, tai pagal XXIX formulą išeis $\sigma = 5$, nes:

$$\sigma = \sqrt{npq} = \sqrt{100 \cdot 0{,}5 \cdot 0{,}5} = \sqrt{25} = 5.$$

Bet jeigu paimsime tos formulos variantą XXIXa, tai

$\sigma = 0{,}05$, nes

$$\sigma = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0{,}5 \cdot 0{,}5}{100}} = \sqrt{\frac{0{,}25}{100}} = \sqrt{0{,}0025} = 0{,}05$$

Panašiai bus ir prie n = 10 000 ar dar didesnio skaičiaus. Pav., jei n = 10 000, tai pagal XXIX formulą išeina $\sigma = \sqrt{npq} = 50$, o formula XXIXa iš karto duoda $\sigma = \sqrt{\frac{pq}{n}} = 0{,}005$, t. y. penkis promilius. Vadinasi, nuo „normos“ arba gausingiausiojo dydžių dalinio normalinis nukrypimas tesudaro tik $\frac{5}{1000}$ arba 0,5%. Šis iš karto santykiniu skaičium pareikštas vid. kvadratingasis nukrypimas ypač pravartus, kai n yra kuris nors daugiaženklis ar įvairiaženklis skaičius (pav. n = 17 385 arba n = 7 436 785) ir nelengva susivokti, kuriame santykyje su visuma yra absoliučiu skaičium pareikštas σ.

Aplamai, čia patiektieji tikimybinių dydžių kvadratingojo nukrypimo pavyzdžiai paryškina didžiųjų skaičių dėsnio veikimą. Vidurinis kvadratingasis nukrypimas prie tos pačios dis-

persijos pobūdžio santykyj su „norma“ arba vidurkiu išeina juo mažesnis, kuo n, t. y. potyrių arba stebėjimų skaičius, didesnis. Binominėj eilutėj, t. y. prie griežtai normalios dispersijos, kai $n = 10$, dar duoda σ, kuris sudaro beveik trečdalį (31,4%) dažniausiojo dydžio $\left(\frac{1,57 \cdot 100}{5} = 31,4\%\right)$. Bet jeigu $n = 100$, tai vidurinis kvadr. nukrypimas tesudaro tik 5%, prie $n = 1\,000$ — 3,1%, $n = 10\,000$ — tik 0,5% ir t. t. Vadinasi, statistikos praktikoj, kai stebimoji masė labai didelė, empiriško stebėjimo keliu gautasis vidurkis net ir iš tokios dalies gali būti laikomas visos masės gana patikimu reprezentantu. Juo labiau, kad vidurinis kvadratingasis nukrypimas (σ), kaip dispersijos rodiklis ir event. maksimalinės tikimybės P_0 matas, galioja ir tiems atvejams, kai $p \neq q$, t. y., kai dviejų nesutaikomų įvykių tikimybiniai dydžiai visumoje pasiskaidę nevienodai. Nes patsai vidurinis kvadratingasis nukrypimas iš tikimybinių dydžių pareina ne tiek nuo dviejų ar kelių nesutaikomų įvykių santykio, kiek nuo *n* skaičiaus. Imkime, pav., dvi mases — a ir b, sutarę, kad abiejose masėse kuri nors žymė sutinkama tuo pat santykiu su priešinga jai žyme, pav., 1 : 9. Vadinasi, abiejose masėse ieškomoji tikimybė $p = 0,1$, ir $q = 0,9$, taigi ir dispersija anormali (atitinka J tipą). Jei tarsime, kad masė a yra atstovaujama 100 vienetų, taigi $n_a = 100$, o $n_b = 10\,000$, tai vidurinis kvadratingasis nukrypimas σ_a masėje a sudarys 3%, o masėje b tik 0,3%. Nes kai $p = 0,1$ ir $n = 100$, tai

$$\sigma_a = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0,1 \cdot 0,9}{100}} = \sqrt{\frac{0,09}{100}} = 0,03.$$

O kai $n_b = 10\,000$, tai

$$\sigma_b = \sqrt{\frac{0,1 \cdot 0,9}{10\,000}} = \sqrt{\frac{0,09}{10\,000}} = 0,003.$$

Savaime suprantama, kad jeigu masę sudaro milijonai vienetų, o jos dalinių savitarpio santykis palieka tas pats, tai vidurinis kvadratingasis nukrypimas bus dar mažesnis. O tuo pačiu pakankamai didelio reprezentantų skaičiaus vaizduojamos „normos“ turi būti dar patikimesnės.

Ši kvadratingojo nukrypimo ypatybė padaro jį tinkama mato priemone ir patikimu rodikliu ne tik st. eilutės vidurkiui arba vad. normaliniam dydžiui (dominantui) charakterizuoti. Imkime, pav., st. eilutę iš variantų $x_1, x_2, x_3, \ldots, x_n$, kurių kiekvienas pasikartoja $y_1, y_2, y_3, \ldots, y_n$ kartų.

Jeigu sutarsime y_i sumą Σy_i žymėti N, o eilutės variantų x_i skaičių masėje žymėsime mažąja n, tai santykio dalinys (d_i) parodys x-inių variantų skaičiaus n_i tikimybinį dydį (p) visoje masėje N. Tuo būdu galime rašyti lygybes:

$$d_{x(i)} = \frac{n_i}{N} = \frac{y_i}{\Sigma y_i}, \text{ arba,}$$

$$p = d_i = \frac{n_i}{N} = \frac{y_i}{\Sigma y_i}.$$

Vadinasi, jeigu turime masę, kur x_i variantų lyginamasis svoris yra atstovaujamas koeficientų arba dalinių $d_1, d_2, d_3, \ldots, d_n$, tai visais atvejais tų dalinių suma $d_1 + d_2 + d_3 + \ldots + d_n = \Sigma d_i = 1$. O jeigu $d_i = p$ ir $\Sigma d_i = 1$, tai, kaip ir anksčiau psl. 251, $q = 1 - p = (d_1 + d_2 + d_3 + \ldots + d_n) - d_i$. Vadinasi, jeigu ieškomoji tikimybė $p = d_2$, tai $q = (d_1 + d_2 + d_3 + \ldots + d_n) - d_2 = \Sigma d_i - d_2$.

O jeigu tarsime, kad dalinių visumoje (Σd_i) dalinys $d_2 = 0{,}3$, tai

$$q = 1 - p = \Sigma d_i - d_2$$
$$= 1 - 0{,}3 = 0{,}7.$$

Pavyzdiniam skaičiavimui mes galime paimti jau naudotus kviečių derliaus duomenis 23 apskrityse, kur parodytas kiekvienos apskrities vidutinis derlius nuo 10,5 ligi 18,2 kvintalų iš 1 ha, o pasėlių plotas sudarė 162,6 tūkst. ha. Lent. 24 tie patys duomenys susaistyti 1 ha intervalais į 9 grupes su x_i pažymiais nuo 10 ligi 18 ha.

Vidutinė kiekvieno iš 23 apskričių derliaus varijanto tikimybė p (=) 0,04, $\left(\frac{1}{23} (=) 0{,}04\right)$.

Lentelė 31.

ŽIEMINIŲ KVIEČIŲ DERLIAI 1922—34 MT.

Eil. (n)	Apskritys	Derliai 100 kg iš 1 ha (x_i)	Plotas 1000 ha (y_i)	Bendras derlius 1000 kvintalų (x_iy_i)
0	1	2	3	4
1.	Telšiai	10,5	2,2	23,10
2.	Utena	11,0	4,6	50,60
3.	Seinai	11,2	2,0	22,40
4.	Trakai	11,3	2,4	27,12
5.	Zarasai	11,4	1,3	14,82
6.	Ukmergė	11,5	6,8	78,20
7.	Kaunas	11,6	10,8	125,28
8.	Alytus	11,9	7,8	92,82
9.	Kretinga	11,9	2,6	30,94
10.	Šakiai	12,6	8,5	107,10
11.	Panevėžys	12,7	15,2	193,04
12.	Marijampolė ..	13,0	10,6	137,80
13.	Tauragė	13,3	4,5	59,85
14.	Rokiškis	13,3	4,0	53,20
15.	Kėdainiai	13,3	14,4	191,52
16.	Raseiniai	13,4	10,7	143,38
17.	Mažeikiai	13,6	7,1	96,56
18.	Biržai	13,6	12,8	174,08
19.	Šiauliai	13,7	24,4	334,28
20.	Klaipėda	14,7	1,2	17,64
21.	Vilkaviškis	15,1	8,2	123,82
22.	Šilutė	15,7	0,1	1,57
23.	Pagėgiai	18,2	0,4	7,28
	Visoj Lietuvoj	$\bar{x}=12,95$	$\Sigma y_i=162,6$	$\Sigma(x_iy_i)=2106,40$

O jeigu imti vienodų intervalų vienodas grupes, kurių yra 9, tai tikimybė $p' = \frac{1}{9} (=) 0{,}11$. Maksimalinė tikimybė P_o (pagal formulą XXVIII) pirmuoju atveju

$$P = \frac{1}{\sqrt{2\pi npq}} = \frac{1}{\sqrt{2 \cdot 3{,}14 \cdot 23 \cdot 0{,}04 \cdot 0{,}96}}$$

$$(=) \frac{1}{\sqrt{6{,}28 \cdot 23 \cdot 0{,}038}} (=) \frac{1}{\sqrt{6{,}28 \cdot 0{,}87}}$$

$$(=) \frac{1}{\sqrt{5{,}46}} (=) \frac{1}{2{,}34} (=) 0{,}427.$$

Antruoju atveju ta pati formula duoda:

$$P' = \frac{1}{\sqrt{6{,}28 \cdot 9 \cdot 0{,}11 \cdot 0{,}89}} (=) \frac{1}{\sqrt{6{,}28 \cdot 9 \cdot 0{,}098}}$$

$$(=) \frac{1}{\sqrt{6{,}28 \cdot 0{,}88}} (=) \frac{1}{5{,}53} (=) \frac{1}{2{,}35} (=) 0{,}426.$$

Vidurinis kvadratingasis nukrypimas pirmuoju atveju išeina:

$$\sigma_{p(n)} = \sqrt{npq} = \sqrt{23 \cdot 0{,}4 \cdot 0{,}96} = \sqrt{23 \cdot 0{,}038}$$

$$= \sqrt{0{,}87} (=) 0{,}93;$$

antruoju:

$$\sigma_{p'(n)} = \sqrt{9 \cdot 0{,}11 \cdot 0{,}89} = \sqrt{9 \cdot 0{,}098} =$$

$$(=) \sqrt{0{,}88} (=) 0{,}94.$$

Kaip matome, tiek maksimalinė derliaus varijantų tikimybė, tiek kvadratingojo nukrypimo atitinkami dydžiai išeina beveik tie patys, ar skaičiuojama iš visumos (iš visų 23 apskričių), ar iš 9 grupių reprezentantų (abiejų dydžių skirtumai nėra didesni kaip 1%). Tai yra svarbi aplinkybė, kuri, visų pirma, rodo, kad šiuo atveju atsisakymas nuo derliaus stebėjimų visose apskrityse ir apsirėžimas tik devyniais rajonais rezultatų beveik nepakeistų. Žymesnę pataisą galėtų įnešti tik didesni kiekvieno derliaus varijanto skirtumai — arba labai skirtingi kiekvieno rajono pasėlių plotai. Panašiais atvejais tektų turi-

mus duomenis vienu ar kitu metodu koriguoti (plėsti stebėjimus, atsižvelgti skirtingų dalinių lyginamojo svorio ar kt.). Šioj vietoj iš 31 lentelės turimų duomenų pakanka konstatuoti, kad medianinis dydis, kuris mūsų skaičiavimuose tenka Marijampolės apskr. ir sudaro M = 13 q iš 1 ha, atstovaujamas 10 600 ha iš 162 600 bendrojo ploto. Taigi atsitiktinei normai mūsų paimtas dydis reprezentuoja apie 6,5% viso ploto ir tuo būdu yra didesnis už atsitiktinį arba „vidutinįjį" tikimybės skaičių (4%) daugiau kai trečdaliu. Iš kitos pusės, toj pačioj lentelėj turime vieną variantą (13,7), kuris atstovaujamas net 24 400 ha arba apie 15%. Bet šie skaičiai nedaug ką gali pasakyti. Derlių variantams pasėlių plotai aprėžti apskritimis, o ne geografiškai statistiniais rajonais. Dėl to jau charakteringesni duomenys iš lentelės 27, kur sudaryti statistiniai rajonai, t. y. suvienodinti derliaus variantai, pagal kuriuos ir išėjo 9 statistiniai rajonai. Čia, kaip matome, didžiausią grupę sudaro 13—14 q, arba, vidutiniškai, 13,5 q derlius iš 1 ha. O ha skaičius siekia 88,5 ir sudaro beveik 55% viso ploto. Šis koeficientas jau žymiai didesnis už maksimalinės „tikimybės" dydį (0,43), nors tikrasis derliaus vidurkis ne visai atitinka 13,5 q normą. Berods, vidurinis kvadratingasis nukrypimas $\sigma = 0{,}94$ q; taigi vidurkio net normalinis patikimumas tėra $13{,}5 \pm 0{,}94 = 12{,}56 - 14{,}44$ q ribose. Iš lentelės 31 mes žinome, kad aritmetinis vidurkis siekia 12,98, o svertinis aritmetinis — 12,95 q. Tuo būdu gautieji vidurkių skirtumai yra vad. normalios arba vidutinės paklaidos ribose.

8. Patikimoji, vidutinė (normali) ir galimoji paklaida statistikos skaičių eilutėse.

Yra vad. statistikos paklaidų, kurios pareina nuo statistinio stebėjimo organizacijos ar technikos trūkumų ir kurių pašalinimas tegalimas tik tuos trūkumus pašalinant. Bet statistikos skaičiai ar jų eilutės gali turėti paklaidų, kurios gali būti susietos su statistinės klasifikacijos bei statistinio analizio metodais ar statistinės technikos priemonėmis. Kai kurios iš tokių paklaidų gali būti ignoruojamos ir jų griežtesnis apibrėžimas nereikalingas. Pav., kiekybinės klasifikacijos ar ir st. analizio darbuose dažnai tenkinamasi artutiniais dydžiais, t. y.

tik riboto griežtumo (apytikriais) skaičiais. Kiekvienas „artutinis“ skaičius slepia savyje mažesnę ar didesnę paklaidą. Tačiau taip pat kiekvienu atveju jau patsai toks artutinis skaičius parodo galimos paklaidos ribas. Pav., jei milijoniniai skaičiai patiekiami su griežtumu ligi vienos dešimtainės ir rašoma, tarsime, 35,0 mil. užuot rašyti griežtai — 35 049 999 ar 34 950 000, tai kiekvienas žino, jog maksimalinė paklaida negali prašokti 100 tūkst. vienetų. O kai iš skaičių matyti, kad paklaida gali būti, ir net žinoma galimos paklaidos ribos, tai statistikai tokia paklaida nebepavojinga.

Kitaip, tačiau, tenka vertinti skaičiavimo bei statistinio vertinimo paklaidas. Pav., mūsų pavyzdžiuose kviečių derliaus tikimybinių dydžių skaičiavimuose gautieji vidurkiai ir kvadratingieji nukrypimai kad ir nedaug kuo vienas nuo kito skiriasi, bet vis dėlto pakankamai, kad galutiniuose derliaus skaičiavimuose sudarytų keliasdešimties ar net 100 tūkstančių centnerių skirtumą. Dėl to ypač visais tais atvejais, kai statistikos skaičiai privalo statistinę masę reprezentuoti, reikalinga kiek galint griežčiau apibrėžti ir normalios bei maksimalinės paklaidos ribas.

Aplamai, jau visas st. eilučių vad. dispersijos analizis siekia statistikos eilutės dydžių paklaidos ribas griežčiau apibrėžti. Tai, kas šių dienų statistinio analizio raštų kalboje vienų vadinama trijų sigmų taisykle, normalia dispersija, ir kt., savo laiku Quetelet ir kt. buvo vadinama, o kai kieno dar ir šiandien vadinama, Gausso paklaidos kreive. Matematika ir matematinė statistika šiandien turi visą eilę formulų ir pagal jas išvestų funkcinių dydžių (taip pat atitinkamų lentelių, panašių logaritmų lentelėms), pagal kuriuos įvairiausios rūšies atvejams ir st. eilutėms su įvairaus tipo dispersija yra galima statistinių dydžių tikimybinės paklaidos apibrėžti. Tačiau šie klausimai sudaro atskirą sritį, kurios nuodugnus išnagrinėjimas išeina iš statistinės metodologijos ribų. Vis dėlto tenka bent su pagrindinėmis vad. statistinių paklaidų apibrėžimo bei jų matavimo tezėmis ir atitinkamomis formulomis susipažinti.

Visų pirma tenka pabrėžti, kad jeigu moderniškoj statistikoj vis gausingiau naudojama ypač „suprastintos“ matematinio analizio priemonės, tai tik tuo pagrindu, kad statistika iš-

imtinai tyrinėja mases arba masinius reiškinius, kur didžiųjų skaičių dėsnio veikimas, paprastai, nepasireikšti negali. O matematikos teikiamieji ir plačiausiai vartojamieji artutiniai bei „suprastinti“ analizio metodai iš begalinio analizio kaip tik yra paremti ta aplinkybe, kad tam tikri vad. begalinio analizio paklaidas apibrėžiantys dydžiai, gaunami gana komplikuoto ir ilgo skaičiavimo rezultate, praktiškai virsta konstantomis, jei tik potyrių skaičiaus n pakankamai didelis. O statistikoj paprastai net ir siauro masto reprezentaciniam stebėjime n būna pakankamai didelis skaičius. Pav., vad. patikimosios paklaidos konstanta 0,67449.... kuri sudaro labai ilgą trupmeną, prie normalios dispersijos pilnumoj galioja jau eilutei iš 30 = n, o praktiškai statistikos dydžiams jau gana patikima ir eilutei iš 9 = n. Tuo tarpu šita konstanta gaunama pagal gana komplikuotą formulą, kur daugiklį sudaro kvadratinė šaknis iš dviejų ($\sqrt{2}$), o dauginamąjį begalinė trupmena 0,476936276..., gaunama iš gana komplikuotos formulos.

Iš kitos pusės, jei imti bet kurią statistinę masę, kur n skaičius labai didelis, tai net prie anormalios dispersijos tokios masės daliniuose artimųjų ir artimiausiųjų varijantų apie bet kurį atskirai paimtą dydį dispersija praktiškai visuomet artima vad. normaliai dispersijai. Pav., visi mes žinome, kad mirusiųjų amžiaus dispersija anormali (priklauso tipo u, nes daugiausia miršta seniai ir maži vaikai). Tačiau, kai kiekvienais metais miršta dešimtys tūkstančių, tai gretimųjų amžiaus grupių mirtingumo kitimai tiek nežymūs ir skirtumai taip nedideli, kad, suskaidžius smulkesniais amžiaus intervalais, beveik kiekvienam atskirai paimtam vienetui iš abiejų pusių bus, palyginti, vienodas (arba nedaug kuo besiskiriantis) artimo amžiaus „draugų“ skaičius. Tą patį arba panašų reiškinį galima sutikti ir kitose vad. statistinėse masėse. O kai taip yra, tai juo didesniu pagrindu galima naudoti ir statistikos paklaidoms apibrėžti formulas bei konstantas, pasiskolintas iš matematinio begalinio analizio arsenalo.

Viena tokia konstanta — tai jau minėtoji vad. patikimosios paklaidos konstanta, kuri statistikos reikalui bus dar gana patikima ir apibrėžus artutine trupmena

0,6745.

Šita konstanta statistikoj naudojama ir naudotina kartu su viduriniu kvadratinguoju nukrypimu vad. patikimajai paklaidai bei dispersijos pobūdžiui griežčiau apibrėžti. Mat, kad kurios nors statistinės eilutės dispersiją būtų galima laikyti normalia, vad. trijų sigmų taisyklė dar reikalinga papildyti patikimosios paklaidos konstantos apibrėžiama taisykle. Pagal šią papildomą taisyklę, 50% visų statistinės eilutės dydžių turi sutilpti intervale, kurį apibrėžia apie a. vidurkį vidurinis kvadratingasis nukrypimas, padaugintas patikimosios paklaidos konstanta pagal formulą:

$$\bar{x} \pm 0{,}6745\sigma \qquad \text{(XXX)}.$$

Vadinasi, griežtai normalią st. eilutės dispersiją charakterizuoja, visų pirma, a. vidurkio ($\bar{x}$), mediano (M_x) ir dominanto (D_x) sutapimas, be to, toksai st. eilutės dydžių pasiskaidymas apie vidurkį, kur sueina:

$\bar{x} \pm 0{,}6745\sigma$	intervalo	ribose	apie	50,00%	visų	dydžių
$\bar{x} \pm 1\sigma$	„	„	„	68,27%	„	„
$\bar{x} \pm 2\sigma$	„	„	„	95,45%	„	„
$\bar{x} \pm 3\sigma$	„	„	„	99,73%	„	„
$\bar{x} \pm 4\sigma$	„	„	„	99,99%	„	„

Tuo būdu vidurinis kvadratingasis nukrypimas tampa rodikliu st. eilutės dispersijos ir jos variantų „patikimumui“, t. y. event. „nukrypimams“ arba „paklaidoms“ (nukrypimams nuo vidurkio ir normalios dispersijos) matuoti. Šias st. eilučių dispersijos apie vidurkį „paklaidas“ klasifikuojant galima paklaidas šitaip apibrėžti:

1. Patikimoji nuo „normos“ (vidurkio) nukrypimų paklaida apibrėžiama intervalo, kurį sudaro vidurinio kvadratingojo nukrypimo (σ) ir patikimosios paklaidos konstantos (0,6745) sandauga pridėjus ir atėmus nuo vidurkio

$$\bar{x} \pm 0{,}6745\sigma$$

Tai reiškia, visų pirma, kad yra 0,5 arba 50% tikimybė, jog iš st. masės ar x-inės eilutės atsitiktinai sutiktas ar paimtas variantas (arba dydis x_i) bus nemažesnis kaip

$\bar{x} - 0{,}6745\sigma$ ir nedidesnis kaip $\bar{x} + 0{,}6745\sigma$. Tokio pat dydžio tikimybė, tačiau, bus ir už tai, kad atsitiktinai paimtas x_i variantas bus mažesnis arba didesnis, negu $\bar{x} \pm 0{,}6745\sigma$ intervalo ribos aprėžia (vadinasi, pateko iš už šio intervalo ribų). Aplamai, tai reiškia, kad 50% visų st. eilutės dydžių neišeina iš $\bar{x} \pm 0{,}6745\sigma$ intervalo ribų, o antra pusė (50%) — yra už šios rūšies ribų;

2. Vidutinė arba normali paklaida x-inių dydžių pasiskaidyme aprėžiama ribų, kurią sudaro $\bar{x} \pm 1\sigma$. Prie normaliai artutinės dispersijos, nemažiau kaip ⅔ visų st. eilutės dydžių (prie griežtai normalios dispersijos — 68,3%) sutelpa šiuo būdu apibrėžto intervalo ribose ir nedaugiau kaip ⅓ (31,7%) — už šių ribų. Iš tokios st. masės atsitiktinai paimto x_i dydžio tikimybė priklausyti dydžių grupės $\bar{x} \pm 1\sigma$ ribose yra lygi 68,3%, o nepriklausyti — 31,7%;

3. Galimoji paklaida apynormalioj statistinės eilutės dispersijoj praktiškai aprėžiama $\bar{x} \pm 3\sigma$ ribos apie vidurkį. Norint, tačiau, tikimybę dar daugiau priartinti tikrenybei, st. praktikoj kai kada galimoji paklaida apibrėžiama ir keturių sigmų apie vidurkį intervalo ribose. Tai reiškia, kad visi statistinės eilutės dydžiai (variantai) yra pasiskaidę trijų (maximum keturių) sigmų intervalo ribose ir praktiškai už šių ribų (t. y. nei mažesnio nei didesnio) neturėtų būti nė vieno x-inės eilutės dydžio.

Visi šie dėsniai galioja vad. normaliajai bei apynormaliajai dispersijai. Pagrindinis statistikos paklaidoms apibrėžti matas, kaip matome, yra vidurinis kvadratingasis nukrypimas. Bet kai vad. kvadratingajam nukrypimui toks svarbus ir atsakomingas vaidmuo tenka statistinio analizio darbų stadijoj, tai juo labiau yra reikalinga, kad patsai kvadratingasis nukrypimas išeitų kiek galint griežtesnis bei patikimesnis dydis.

Kuria prasme ir kuriuo būdu šitą vid. kvadratingojo nukrypimo, arba sigmos, patikimumą galima tikrinti ir patikrinti, t. y. kas gali nulemti jo nepatikimumą?

Visų pirma, tenka turėti galvoje, kad vidurinis kvadratingasis nukrypimas, vienu ar kitu metodu skaičiuojamas, visuo-

met yra funkcinėj pareinamybėj su vidurkiu. Vadinasi, jei gautasis arba naudotasis vidurkis dėl vienų ar kitų priežasčių bus klaidingas ar nepatikimas, tai tuo pačiu bus nulemtas ir vid. kvadratingojo nukrypimo nepatikimumas. Dėl to pirmiausia tenka patikrinti vidurkio patikimumą. Šis patikimumas, visų pirma, matuojamas vidurkio patikimąja paklaida, kuri apibrėžiama skaičiumi, gautu iš kvadratingojo nukrypimo su patikimosios paklaidos konstanta sandaugos, padalintos kvadratine šaknimi iš eilutės narių skaičiaus. Vadinasi, paties vidurkio patikrinime ligi tam tikro laipsnio dalyvauja ta pati sigma (vid. kvadratingasis nukrypimas). Pravartu, todėl, pirmiausia patikrinti vid. kvadratingasis nukrypimas. Jei sutarsime bet kurio statistinio dydžio patikimąją paklaidą žymėti raide *u*, tai kvadratingojo nukrypimo patikimoji paklaida bus galima apibrėžti pagal formulą:

$$u_\sigma = \pm \frac{0{,}6745\sigma}{\sqrt{2n}} \qquad \text{(XXXI).}$$

Kvadratingojo nukrypimo patikimoji paklaida pirmoj eilėj nulemiama n dydžio: juo didesnis st. eilutės narių (tyrinėjimų vienetų) skaičius, tuo mažesnė, prie lygių sąlygų, turės išeiti patikimoji paklaida, ir atvirkščiai. Paties aritmetinio vidurkio patikimoji paklaida *u* gaunama pagal formulą:

$$u_{\bar{x}} = \frac{0{,}6745\sigma}{\sqrt{n}} \qquad \text{(XXXII).}$$

Patikimosios paklaidos u matematiškai statistinė prasmė yra ta, kad tikimybiniu atžvilgiu duotasis statistinis dydis prie normalios dispersijos gali būti mažesnis arba didesnis tiek, kiek sudaro prie jo pridėjus ir atėmus pagal atitinkamą formulą gautąją vad. patikimosios paklaidos skaičių u.

Pavyzdiniam skaičiavimui paimkime tuos pačius, mūsų jau nekartą naudotus, kviečių derliaus duomenis. Kai svertinis aritmetinis vidurkis $\bar{x} = 12{,}95$, kvadratingasis nukrypimas $\sigma = 1{,}1$, tai kvadratingojo ir vidurkio nukrypimo patikimosios paklaidos pagal atitinkamas formulas, — jei tarti, kad duomenys gauti tobulai organizuoto ištisinio stebėjimo keliu ir $n = 162{,}6$, gaunamas:

1. $u_\sigma = \frac{0{,}6745 \cdot \sigma}{\sqrt{2n}} = \frac{0{,}6745 \cdot 1{,}1}{\sqrt{2 \cdot 162{,}6}} = \frac{0{,}74195}{\sqrt{325{,}2}}$ (=)

(=) $\frac{0{,}74195}{18{,}033}$ (=) 0,04.

2. $u_{\bar{x}} = \frac{0{,}6745 \cdot \sigma}{\sqrt{n}} = \frac{0{,}6745 \cdot 1{,}1}{\sqrt{162{,}6}}$ (=) $\frac{0{,}74195}{12{,}75}$ (=) 0,06.

Kadangi mūsų naudojamame pavyzdyje x-inės eilutės dydžiai (derliai) buvo žymimi dvigubais centneriais — kvintalais, t. y. 100 kg, o n = 162,6 atstovavo tūkstančius ha, vadinasi, sveikais racionaliais skaičiais n = 162 600, — tai patikimosios paklaidos gautieji skaičiai tenka dar padalyti 1 000, ir tuo būdu gauname:

$$u_\sigma = \frac{0{,}04}{1000} = 0{,}00004 \text{ (q) ir}$$

$$u_{\bar{x}} = \frac{0{,}06}{1000} = 0{,}00006 \text{ (q).}$$

Norėdami aiškiai šių skaičių statistinę prasmę pažinti, paverskime x-inės eilutės dydžius sveikais skaičiais (kilogramais), $\bar{x}$, σ, u daugindami 100. Tuo būdu gausime:
$\bar{x} = 12{,}95 \cdot 100 = 1\,295$ (kg) ir $u_{\bar{x}} = 0{,}00006 \cdot 100 = 0{,}006$ (kg)
$\sigma = 1{,}1 \cdot 100 = 110$ (kg) ir $u_\sigma = 0{,}00004 \cdot 100 = 0{,}004$ (kg)

Vadinasi, vidurkio patikimoji paklaida tesudaro tik $\frac{6}{1000}$ kg arba 6 gramus, o vidurinio kvadratingojo nukrypimo patikimoji paklaida tesudaro tik $\frac{4}{1000}$ kg arba 4 gramus. Taigi vidurkis negali būti mažesnis kaip 1294,994 kg nė didesnis kaip 1295,006 kg; v. kvadratingasis gi nukrypimas negali būti mažesnis 109,6 kg, nė didesnis kaip 110,4 kg. O grįžtant prie pirmykščių skaičių, patikimosios paklaidos apibrėžiamos ribos sudaro:

$\bar{x} \pm u_{\bar{x}} = 12{,}95 \pm 0{,}00006$ arba 12,94994 — 12,95006 (q).

$\sigma \pm u_\sigma = 1{,}10 \pm 0{,}00004$ arba 1,09996 — 1,10004 (q).

Kaip matome, patikimoji paklaida abiejais atvejais yra tiek maža, kad statistikos praktikoj į tokią paklaidą galima visai dėmesio nekreipti. Bet patikimosios paklaidos reikšmė tuojau ir gana žymiai padidėja, jei tarsime, kad derliaus vidurkis $\bar{x}$ ir jį atitinkantis kvadr. nukrypimas buvo gautas ne ištisiniu stebėjimu viso pasėlių ploto, kurį sudarė 162 600 ha, o reprezentaciniu metodu, — sakysim, buvo stebimas tik 1 ha iš tūkstančio. Jei tarsime, kad kiekvienoj apskrityje ištyrus tiek ha, kiek užsėta tūkstančių ha, būtų gauti tie patys rezultatai, tai aritmetinis vidurkis ($\bar{x}$) ir kvadr. nukrypimas (σ) išeina tie patys (12,95 ir 1,1 q), bet tų dydžių patikimoji paklaida, kai n = 162,6, o ne 162 600, atitinkamai sumažėja ir gaunama: $u_{\bar{x}} = 0{,}06$, o $u_{\sigma} = 0{,}04$ q. Tada jau vidurkis patikimosios paklaidos ribose išeina:

$$\bar{x} \pm u_{\bar{x}} = 12{,}95 \pm 0{,}06 \text{ arba } 12{,}89 — 13{,}01 \text{ (q) iš 1 ha,}$$

$$\sigma \pm u_{\sigma} = 1{,}10 \pm 0{,}04 \text{ arba } 1{,}06 — 1{,}14 \text{ (q) iš 1 ha.}$$

Skirtumas, kaip matome, atrodo nemažas. Bet dar didesnis bus skirtumas, jei tarsime, kad kiekvienoj apskrityje, nepriklausomai nuo pasėlių ploto, buvo stebimas tik 1 ha derlius ir tuo pačiu laiku sutarsime, kad rastieji derliai atitinka mūsų naudotoj lentelėj patiektuosius. Tokiu atveju mes galime pasikliauti paprastuoju aritmetiniu vidurkiu, kuris sudaro $\bar{x} = 12{,}98$ q, arba medianu, kur M = 13,00 q. Skirtumas, kaip matome, visai nedidelis.

Tačiau vidurinis kvadratingasis nukrypimas čia gaunamas žymiai didesnis — net apie 1,72:

$$\sigma = \sqrt{\frac{x_2^2}{n} - \bar{x}^2} = \sqrt{\frac{3943{,}19}{23} - 12{,}98^2} (=)$$

$$(=) \sqrt{171{,}44 - 168{,}48} (=) \sqrt{2{,}96} (=) 1{,}72.$$

Atitinkamai, žinoma, pasikeičia ir patikimoji paklaida. Pagal atitinkamas formulas, patikimosios paklaidos išeina:

$$u_{\bar{x}} = \frac{0{,}6745 \cdot \sigma}{\sqrt{n}} = \frac{0{,}6745 \cdot 1{,}72}{\sqrt{23}} (=) \frac{1{,}16}{4{,}80} (=) 0{,}24.$$

$$u_\sigma = \frac{0{,}6745 \,.\, 1{,}72}{\sqrt{46}} = \frac{1{,}16}{6{,}78} (=) \; 0{,}17.$$

Vadinasi, atitinkamų dydžių patikimumas aprėžiamas:

$\bar{x} \pm u_{\bar{x}} = 12{,}98 \pm 0{,}24$, t. y. 12,74 — 13,22 (q),

$\sigma \pm u_\sigma = 1{,}72 \pm 0{,}17$, t. y. 1,55 — 1,89 (q).

Kaip matome, gali susidaryti nemažas skirtumas, jei iš kiekvieno ha derliaus išeina apie 24 kg, t. y. beveik ½ centnerio, mažesnis arba didesnis, negu numatyta.

Tokiose aplinkybėse, kai reprezentantai ne vienodo dydžio st. masės daliniuose pasiskaidę (mūsų pavyzdyje atskirose apskrityse nevienodi pasėlių plotai), o aritmetinis vidurkis ir medianas maža kuo tesiskiria, tai yra labai didelė tikimybė, kad visoje masėje dydžių dispersija artima normaliai ir tuo būdu statistiškai prasmingesni bei patikimesni gali būti tikimybiniai dydžiai. Iš tikrųjų, net turint tik a. vidurkį $\bar{x} = 12{,}98$, $M = 13{,}00$ ir $\sigma = 1{,}72$, asimetrijos koeficientas a_σ išeina artimas 0, nes:

$$a_\sigma = \frac{3\,(\bar{x} - M)}{\sigma} = \frac{3\,(12{,}98 - 13{,}00)}{1{,}72} = \frac{-\,0{,}02 \,.\, 3}{1{,}72}$$

$$= \frac{-\,0{,}06}{1{,}72} (=) \; -\,0{,}03.$$

Tuo tarpu mes žinome, kad normalios dispersijos dėsningumais paremti tikimybiniai dydžiai savo statistinės prasmės pradeda nustoti tik tada, kai asimetrijos koeficientas $a_\sigma > 0{,}5$ t. y. artinasi 1.

Šį patikrinimą padarę, galime panaudoti dydžius iš psl. 256 Kai $n = 23$, tai bet kurio iš 23 varijantų x_i tikimybė $p_i = \frac{1}{23} (=) \; 0{,}04$. Tuo būdu $q = 1 - 0{,}04 = 0{,}96$, ir dėl to σ pagal XXIX formulą

$$\sigma = \sqrt{23 \,.\, 0{,}04 \,.\, 0{,}96} (=) \; 0{,}93.$$

Čia jau gautas vid. kvadratingasis nukrypimas, kuris visai artimas σ, gautam mūsų tariamame ištisinio stebėjimo skaičiavime. Savaime suprantama, kad patikimoji paklaida, kai

n = 23, — o ne 162 600 ir net ne 162,6, — turi išeiti didesnė. Atitinkamų formulų panaudojimas duoda šitokias patikimąsias paklaidas:

$$u_{\bar{x}} = \frac{0{,}6745\sigma}{\sqrt{n}} = \frac{0{,}6745 \cdot 0{,}93}{\sqrt{23}} = \frac{0{,}62785}{4{,}7958} \ (=)\ 0{,}13\ (q).$$

$$u_{\sigma} = \frac{0{,}6745 \cdot 0{,}93}{\sqrt{2 \cdot 23}} = \frac{0{,}62785}{6{,}7823} (=)\ 0{,}0896\ (=)\ 0{,}09\ (q).$$

Tuo būdu vidurinio kvadratingojo nukrypimo ir aritmetinio vidurkio, sutarus, kad $\bar{x} = M$; patikimoji paklaida aprėžiama:

$\bar{x} \pm u_{\bar{x}} = 12{,}98 \pm 0{,}13$, t. y. 12,85 — 13,11 (q), o

$\sigma \pm u_{\sigma} = 0{,}93 \pm 0{,}09$ arba 0,84 — 1,02 (q).

Tuo būdu galime įsitikinti, kad matematinės tikimybių teorijos dėsniai bei metodai statistikos praktikai tam tikromis sąlygomis duoda skaičius, kurie labai artimi tikrenybei. Berods, galima dar prikišti, kad pačių tikimybinių dydžių patikimosios paklaidos ribos nežinomos. Bet mes jau matėme (psl. 252-4), kad tikimybiniai dydžiai pareina ne tiek nuo p : q santykio, kiek nuo n. Be to, tikimybinių dydžių p_i arba dalinio n_i paklaidos ribos taip pat aprėžiamos patikimąja paklaida pagal formulą:

$$u_{p_i} = 0{,}6745 \sqrt{\frac{pq}{n}}, \qquad \text{(XXXIII).}$$

arba

$$u_{n_i} = 06745 \sqrt{npq}, \qquad \text{(XXXIIIa).}$$

jeigu norima aprėžti patikimąją paklaidą tikimybei p_i atstovaujantį absoliutų skaičių n_i.

Mūsų pavyzdyje:

$$u_{p_i} = 0{,}6745 \sqrt{\frac{pq}{n}} = 0{,}6745 \sqrt{\frac{0{,}04 \cdot 0{,}94}{23}} \ (=)$$

$$(=)\ 0{,}6745 \sqrt{\frac{0{,}038}{23}} \ (=)\ 0{,}6745 \sqrt{0{,}0017} \ (=)$$

$$(=)\ 0{,}6745 \cdot 0{,}0412\ (=)\ 0{,}028.$$

Kai $p = 0{,}04$, o $u_p = 0{,}028$, tai tenka konstatuoti, jog p_i patikimoji paklaida gana didelė, nes jos intervalas

$$p_i \pm u_p = 0{,}04 \pm 0{,}028;$$

taigi galutinai p_i yra 0,012 — 0,068 ribose.

Bet patikrinkime, kiek skirtųsi $\sigma_{p,n}$, jei tikrai $p = 0{,}012$ arba 0,068.

Jeigu $p = 0{,}012$, tai $q = 0{,}988$ ir tuo būdu

$$\sigma = \sqrt{npq} = \sqrt{23 \cdot 0{,}012 \cdot 0{,}988} \,(=)\, \sqrt{23 \cdot 0{,}012} =$$
$$= \sqrt{0{,}276} \,(=)\, 0{,}524.$$

Jeigu $p = 0{,}068$, tai $q = 0{,}932$ ir

$$\sigma = \sqrt{23 \cdot 0{,}068 \cdot 0{,}932} \,(=)\, \sqrt{23 \cdot 0{,}063}$$
$$(=)\, \sqrt{1{,}45} \,(=)\, 1{,}2.$$

Kaip matome, net blogiausiu atveju tikimybinio skaičiavimo keliu gautasis vidurinis kvadratingasis nukrypimas nuo „tikrojo", t. y. patikimiausiojo, nukrypsta kiek daugiau kaip 40%.

Šiais pavyzdžiais pakankamai įrodomos tikimybių teorijos bei atitinkamo matematinio analizio metodų svarba ir nauda statistikai. Juo labiau, kad statistikos praktikoj tokių kraštutinių atsitikimų, kur tektų operuoti tikimybiniais dydžiais $p < 0{,}1$, t. y. mažesniais negu 10%, retai tegali pasitaikyti. O jeigu p nėra mažesnis kaip 0,1 ir potyrių arba stebėjimų skaičius nemažesnis kaip $n = 23$, tai

$$\sigma = \sqrt{npq} = \sqrt{23 \cdot 0{,}1 \cdot 0{,}9} = \sqrt{23 \cdot 0{,}09} = \sqrt{2{,}07} \,(=)\, 1{,}44.$$

Tikimybinio dydžio patikimoji paklaida tokiu atveju išeitų:

$$u_p = 0{,}6745 \sqrt{\frac{0{,}1 \cdot 0{,}9}{23}} \,(=)\, 0{,}6745 \sqrt{0{,}004} \,(=)\, 0{,}06.$$

Tuo būdu $p_i \,(=)\, 0{,}04$ išeitų tik blogiausiu, vadinasi, mažiausia patikimu, atveju.

Šie pavyzdžiai vis dėlto parodo, kad viena patikimoji paklaida u bet kurio statistinio dydžio galimiems nuo normos nukrypimams apibrėžti nepakankama, o, panašiai kaip st. eilutės dydžių dispersijos galimiems nukrypimams apibrėžti, pravartu dar vidutinės arba normalios, o taip pat ir g a l i m o s paklaidos ribas žinoti. Paprastai tokiais atvejais pasikliaujama vad. trijų sigmų taisykle, nors šiuo atveju turima reikalas ne su σ (viduriniu kvadratinguoju nukrypimu), o su u, kuri atstovauja tik pirmąjį galimų paklaidų laipsnį. Antrąjį bet kurio statistinio dydžio paklaidos laipsnį, kaip ir nuo vidurkio galimų nukrypimų apibrėžimuose, apibrėžia $\pm 2u$ (dviejų patikimosios paklaidos dydžių ribos) ir tretįjį — galimos paklaidos ribas — apibrėžia $\pm 3u$.

9. Tikimybės ir atsitiktinės paklaidos reprezentacinėj statistikoj.

a) Bendros pastabos.

Tikimybių skaičiavimo metodai bei tikimybiniai dydžiai statistinę prasmę ir svarbą įgauna ypač reprezentacinėj statistikoj. O reprezentacinis metodas gali būti ir būva naudojamas visose statistikos darbo fazėse, t. y. ne tik stebėjime, bet ir vaizdavime ir, žinoma, statistikos analizy. Pav., gana dažnai ištisinio stebėjimo metodu gautieji duomenys reprezentaciniu metodu tiriami (klasifikuojami bei analizuojami), kad būtų galima pačiu trumpiausiu laiku charakteringas tyrinėjamos masės ypatybes, kad ir apytikriai, pažinti ir jau paskui tas ypatybes atitinkančius statistinės masės dalinius, medžiagą galutinai apdirbant bei klasifikuojant, sudaryti.

Statistiškai metodologiniu atžvilgiu šioj darbo stadijoj (kai reprezentaciniu metodu yra tiriama ištisiniu stebėjimu gautoji medžiaga) visi darbo uždaviniai, taigi ir problemos, yra tos pačios rūšies, kaip ir reprezentaciniu metodu daromame tiesioginiame stebėjime. Vadinasi, ir čia reprezentantai gali ir privalo būti vienu ar kitu metodu parenkami; taip pat ir turimos statistinės medžiagos charakteringųjų ypatybių tyrimas gali tos pačios rūšies bei apimties paklaidų turėti, kaip betarpiškame

stebėjime. Dėl to yra be galo svarbu pasiaiškinti, kokios rūšies ir apimties reprezentacinėj statistikoj paklaidos gali pasidaryti aplamai imant.

Jau II dalyj buvo parodyta ir įrodyta, kad, nekomplikuotą reiškinį stebint reprezentaciniam metodui ideališkomis sąlygomis, gautieji duomenys negali išeiti nei blogesni, nei mažiau patikimi, kaip ištisiniu stebėjimu gautieji. Tiesa, statistikos praktikoj visai ideališkų stebėjimo sąlygų retai pasitaiko. Iš kitos, tačiau, pusės, statistika — ne bugalterija, kur būtų reikalingas ligi cento sąskaitų tikslumas bei griežtumas. Be to, ideališkos reprezentaciniam metodui sąlygos yra visais tais atvejais, kai statistinės masės vienetų visumos, ar bent jų statistiniuose daliniuose, dispersija yra normali — ar bent artutinė normaliai. Kai taip yra, tai reprezentacinis metodas gali ir privalo būti laikomas vienodai patikimu ne tik tais atvejais, kai yra stebimas pakankamai didelis atsitiktinių reprezentantų skaičius, bet ir tuomet, kai yra statistinės masės dalinių tipiškieji reprezentantai. Tik gautųjų duomenų patikimumas, t. y. paklaidų rūšys bei jų tūris (ribos) vienu ir antru atveju privalo šiek tiek skirtingų patikrinamųjų statistinio analizio priemonių bei metodų.

Visų pirma, tenka išaiškinti bei pasiaiškinti paklaidų ribas, pareinančias betarpiškai nuo visos statistinės masės ir reprezentaciniu metodu tyrinėjamų vienetų skaičiaus. Jei tarsime, kad statistikos stebimoji arba analizuojamoji masė dispersijos atžvilgiu yra normali, tai vis dėlto, kaip buvo parodyta, patikimoji statistinių dydžių paklaida gali išeiti perdaug didelė vien tik dėl to, kad stebimų vienetų skaičius yra permažas ir vad. didžiųjų skaičių veikimas, kuris atsitiktines statistinio stebėjimo paklaidas eliminuoja, nepakankamai pasireiškia. Pav., reprezentaciniu metodu stebint javų derlius, galima atsitiktinai pataikinti tokį ūkininką, kur rastasis 1 ha derlius atitiks ideališkąjį viso Lietuvos ploto 1 ha derliaus vidurkį. Tačiau remtis tokio stebėjimo duomenimis vistiek neturėtume teisės: nes tikimybė apsirikti tokiomis sąlygomis mažiausia yra tokia pat, kaip ir tikrenybę (ideališkąjį vidurkį) pataikyti. Nedidelė tikimybė prie normalios dispersijos dar ir 10 atsitiktinių stebėjimų darant: nes, kaip žinoma, tokiomis sąlygomis „normali-

nio“ dydžio, t. y. ideališkojo vidurkio charakterį apibrėžiantis vid. kvadratingasis nukrypimas, skaičiuojamas pagal formulą

$$\sigma = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0{,}5 \cdot 0{,}5}{10}},$$

išeina beveik 16%.

Tuo būdu reprezentacinėj statistikoj galimų paklaidų ribas net ir ideališkiausiomis sąlygomis visų pirma nulemia reprezentantų skaičius, be to, tikslumas arba griežtumas, kurio iš statistikos duomenų reikalaujama. Jei statistikos stebimas reiškinys, arba statistikai statomas uždavinys, yra tokios rūšies, kad, sakysime, 5% paklaida gali būti laikoma nereikšminga, tai sulig tuo, kaip didelė tyrinėjamos masės visuma, gali būti reikalingas mažesnis ar didesnis atsitiktinių reprezentantų skaičius. Bet jeigu iš statistikos duomenų reikalaujama griežtumo ligi 98% (paklaida leidžiama tik ligi 2%), tai reprezentantų reikės daugiau.

Sutarkime st. masės bendrąjį vienetų skaičių žymėti didžiąja N, o jos reprezentantų skaičių n. Jei reprezentaciniu metodu norėsime ištirti tokios masės bet kurių pažymių proporcionalų santykį, tai čia statistikos paklaidai galios jau mums pažįstamos tikimybiniams dydžiams apibrėžti formulos XXVIII variantai ryšium su formulomis III, XXIX ir XXXIII. Pagrindinė sąlyga reprezentaciniam metodui funkcionuoti — tai pakankamai didelė st. masė. Tarkime, kad turime dvi mases N_1 ir N_2. Sutarkime, kad $N_1 = 1\,000$, o $N_2 = 100\,000$. Taip pat sutarkime, kad abiejose masėse mums rūpimas pažymys — sakysime, mokyklinio amžiaus berniukų ir mergaičių proporcingi skaičiai, atstovaujami vienodai, — vadinasi, ir berniukų (mergaičių) daliniai (d_j) abiejose masėse yra vienodi. Jei sutarsime mums rūpimą pažymį turintį masės narių skaičių žymėti raide m, tai atitinkami masės daliniai

$$d_1 = \frac{m}{N_1} = \frac{500}{1\,000} = 0{,}5 \text{ ir}$$

$$d_2 = \frac{m}{N_1} = \frac{50\,000}{100\,000} = 0{,}5.$$

Dabar klausimas: kaip didelis gali būti skirtumas abiejų masių tikimybiniuose dydžiuose (p_j), jei reprezentantų skaičių

paimti vienodai po 100, t. y. jeigu $n_1 = n_2 = 100$. Statistiškai metodologiniu atžvilgiu šis skirtumas yra atitinkamo st. dalinio reprezentacinės tikimybės (p_i) paklaida, kurią galime, sutarę, žymėti raide u ir gauti pagal formulą

$$u = \frac{m}{N} - \frac{m-n}{N-n} \qquad \text{(XXXIV).}$$

Jei mes turime dvi mases iš kurių $N_1 = 1000$, o $N_2 = 100\,000$, ir abiejose masėse $m = 0{,}5$, tai, abiejose masėse stebint po 100 reprezentantų ($n = 100$), gausime:

$$u_1 = \frac{m_1}{N_1} - \frac{m_1 - n_1}{N_1 - n_1} = \frac{500}{1000} - \frac{500-100}{1000-100} =$$

$$= \frac{1}{2} - \frac{4}{9} (=) 0{,}5 - 0{,}444 \ldots (=) 0{,}056.$$

Antruoju atveju, kai $N_2 = 100\,000$, $m_2 = 50\,000$ ir $n_2 = 100$:

$$u_2 = \frac{50\,000}{100\,000} - \frac{50\,000-100}{100\,000-100} = \frac{5}{10} - \frac{500-1}{1000-1} =$$

$$= \frac{1}{2} - \frac{499}{999} (=) 0{,}5000 - 0{,}4995 = 0{,}0005$$

Kaip matome, prie pakankamai didelio stebimos masės vienetų skaičiaus N, pačioje st. masės sudėtyje pirmosios serijos (pirmųjų stebėjimų) reprezentantų net kraštutinis nuo normos nukrypimas tiek maža ką pakeičia, kad labai maža tikimybė, kad ir sekanti reprezentantų serija galėtų išeiti taip pat vienašališka. Be to, kiekvienas sekantis stebėjimas, kai st. masės vienetų skaičius aprėžtas, nuo normos nukrypimų tikimybę sumažina. Pav., jei dėžutėje bus 500 baltų ir 500 raudonų rutuliukų, ir jeigu pirmąjį šimtą traukdami pataikysime tik baltos spalvos rutuliukus, tai liks dėžutėje tik 400 baltųjų ir vien 500 raudonųjų rutuliukų. O tuo pačiu ir stebėjimo sąlygos pasikeičia: jei pirmąjį kartą traukiant, baltą rutuliuką ištraukti tikimybė $p = 0{,}5$, tai šimtą pirmąjį kartą traukiant toji tikimybė bus sumažėjusi ir pasireikš $p = 0{,}444\ldots$, nes:

$$\frac{500-100}{1000-100} = \frac{5-1}{10-1} (=) 0{,}444\ldots$$

Jei atsitiktų, kad ir sekanti traukimų serija duotų tik baltuosius rutuliukus, tai vėlesniame bandyme šansai baltiesiems sumažėtų ligi $0{,}375\left(\frac{400-100}{900-100}=0{,}375\right)$. Ir čia yra įrodymas, kad reprezentaciniu metodu stebint statistinę, t. y. aprėžtą masę, tikimybės bus juo artesnės tikrenybėms, kuo reprezentantų (stebėjimų) skaičius didesnis. Vadinasi, čia ir statistinių tikimybių paklaidos, kai jos matuojamos begalinio skaičiavimo analizio metodais bei priemonėmis, negali būti didesnės kaip matematinių tikimybių paklaidos. O iš to išeina, kad reprezentacinės statistikos praktikoj st. masės stebėjimo bei skaičiavimo paklaidų ribos gali būti siauriau ir griežčiau apibrėžiamos, kaip matematikos tikimybėse.

b) Reprezentacinės statistikos dalinių tikimybė bei paklaidos.

Pagrindinis klausimas, kuris reprezentacinės st. praktikoj visa ką nulemia, tai šis: kaip didelis turi būti atsitiktinių reprezentantų skaičius n iš st. masės N, kad bet kurio dalinio d_i vidutinė paklaida $\bar{u}$ neprašoktų apibrėžtojo dydžio.

Jei sutarsime vidutinę paklaidą žymėti $\bar{u}$, N masėje bet kurio statistinio dalinio tikimybinį dydį p, o atsitiktinių reprezentantų skaičių raide n, tai vidutinė tikimybinio dydžio paklaida $\bar{u}_p$ galima gauti pagal formulą

$$\bar{u}_p=\sqrt{\frac{p(1-p)}{n}\left(\frac{N-n}{N-1}\right)}(=)\sqrt{\frac{p(1-p)}{n}\left(1-\frac{n}{N}\right)}\text{(XXXV).}$$

Ši formula, kaip lengva įsitikinti, duoda ne ką kitą, kaip šiek tiek sumažintą vidurinį kvadratingąjį nukrypimą. Nes jeigu sutarsime, kad p = 0,5, N = 10 000 ir n = 1000, tai

$$\bar{u}_p=\sqrt{\frac{p(1-p)}{n}\left(\frac{N-n}{N-1}\right)}=\sqrt{\frac{0{,}5\,.\,0{,}5}{1000}\left(\frac{10\,000-1000}{10\,000-1}\right)}=$$

$$=\sqrt{\frac{0{,}25}{1000}\left(\frac{9000}{9999}\right)}\;(=)\;\sqrt{0{,}00025\,.\,0{,}9001}\;(=)$$

$$(=)\sqrt{0{,}000225025}\;(=)\;0{,}015.$$

Artutinė paklaida praktiškai ta pati:

$$\bar{u}'_p = \sqrt{\frac{p\,(1-p)}{n}\left(1-\frac{n}{N}\right)} = \sqrt{\frac{0{,}25}{1000}\left(1-\frac{1000}{10\,000}\right)} =$$

$$= \sqrt{0{,}00025\,.\,0{,}9} = \sqrt{0{,}000225} = 0{,}015.$$

Jei skaičiuotume ne u, o σ iš n = 1000, tai gautume pagal formulą XXIXa:

$$\sigma = \sqrt{\frac{p\,(1-p)}{n}} = \sqrt{\frac{0{,}5\,.\,0{,}5}{1000}} = \sqrt{0{,}00025}\ (=)\ 0{,}016.$$

Kaip matome, tarp u_p ir σ_p gautasis skirtumas nedidelis. Jis tik tuo atveju išeitų kiek didesnis, jei n santykyje su N būtų didesnis[8]. Tačiau tuo pačiu pasakoma, kad skirtumas tarp ū ir σ turi būti juo didesnis, kuo n (reprezentantų skaičius) santykyje su N (st. masės vienetų skaičiumi) didesnis. O tatai ir neleidžia šių dviejų st. analizio dydžių sukeisti.

Iš antros pusės, patiektoji XXXV formula reprezentaciniu metodu gautajai st. dalinio vidutinei paklaidai apibrėžti yra neparanki dėl to, kad yra reikalinga st. masės reiškinį atstovaujančio dalinio tikrosios tikimybės (p), kurią apibrėžti tegalima tik žinant reiškinį atstovaujančių x-inės eilutės vienetų skaičių y_i. Nes p, pritaikius formulų III bei XXVIIIa varijantus, gaunama:

$$p = d_i = \frac{x_i}{X} = \frac{n_i}{N} = \frac{y_i}{\Sigma y_i}.$$

Bet, jei bet kurio reiškinio tikroji proporcija N masėje yra jau žinoma, tai nereikalingi tikimybiniai dydžiai, nei reprezentacinis metodas tiems dydžiams rasti. Dėl to ir XXXV formula naudotina tik ypatingais atsitikimais (pav., kai yra pagrindo tarti, kad kurio nors reiškinio santykinis dydis ir reprezentacinis x_i stebimoje nežimoje X masėje yra toks pat, kaip žinomas p dydis ištirtoje N masėje). Aplamai, repre-

---

[8] Jeigu n = 100, tai $\sigma = \sqrt{\frac{0{,}5\,.\,0{,}5}{100}} = 0{,}05$. Tuo tarpu $\bar{u}'_p$ prie to paties N, išeina: $\bar{u}'_p \sqrt{\frac{0{,}25}{100}\left(1-\frac{100}{10\,000}\right)}$ (=) 0,0497.

zentaciniu metodu stebimoje masėje dalinio (d_j) tikimybės vidutinei paklaidai rasti naudojamas XXXV formulos pakeistas variantas, kur p atstovauja p′, kuri gaunama jau betarpiškai iš reprezentaciniu metodu ištirtųjų N masės atsitiktinių vienetų skaičiaus n, o p′ nulemia x_j pažymio reprezentantai $\dot{x}$.

Jei sutarsime, kad $\Sigma x_j = X = N$, $\Sigma \dot{x} = x = n$,

tai

$$p = \frac{x_j}{X} = \frac{n_j}{N},$$

o

$$p' = \frac{\dot{x}}{x} = \frac{n'}{n}.$$

Tuo būdu XXXV formulos variantas a duos:

$$\bar{u}_{p'} = \sqrt{\frac{p'(1-p')}{n}\left(1-\frac{n}{N}\right)} \qquad \text{(XXXVa)}.$$

Sulig tuo, kiek skirsis p′ nuo p, išeis skirtinga ir vidutinė paklaida ū. Tačiau, jeigu tik n (reprezentantų skaičius) nėra labai mažas, tai esminio skirtumo būti negali, nes ū visuomet pareina daugiau nuo n ir nuo n bei N dydžių santykio, negu nuo p (1 — p) arba nuo (p′ : q′). Be to, pravartu įsidėmėti, kad maksimalinis p (1 — p) sandaugos dydis gaunamas binomo paskaidoje, kai $p = 1 - p = q = 0{,}5$. Tuo būdu ir p′ (1 — p′) sandauga niekuomet negali būti didesnė kaip 0,25 arba $\frac{1}{4}$. Dėl šios priežasties vidutinę p tikimybės paklaidą kai kas skaičiuoja iš karto pagal specialinį XXXV formulos variantą, kur

$$\bar{u}_{p'} = \sqrt{\frac{0{,}25}{n}\left(1-\frac{n}{N}\right)} = \sqrt{\frac{1}{4n}\left(1-\frac{n}{N}\right)} \qquad \text{(XXXVb)}.$$

Praktiškai šis variantas turi duoti maksimalinę, t. y. galimąją paklaidą u_p visais tais atvejais, kai p ir q dydžiai žymiau skiriasi, ir visai artimą vidutinei paklaidai ($\bar{u}_p$), kai p ir q yra artimi dydžiai. Pav., kai mūsų pavyzdiniame skaičiavime $p' = 0{,}5$, $n = 1000$ ir $N = 10\,000$, tai formulos XXXVb variantas duoda:

$$\bar{u}_{p'} = \sqrt{\frac{1}{4n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{1}{4 \cdot 1000}\left(1 - \frac{1000}{10\,000}\right)} =$$

$$= \sqrt{\frac{1}{4000} \cdot 0{,}9} = \sqrt{\frac{9}{40\,000}} = \frac{3}{200} = 0{,}015, \text{ t. y.}$$

visai tą patį dydį, kaip ir pagal formulą XXXVa. Tačiau, jei būtų $p' = 0{,}1$, tai formula XXXVb duotų taip pat $\bar{u}_{p'} = 0{,}015$ o formula XXXVa —

$$\bar{u}_{p'} = \sqrt{\frac{0{,}1 \cdot 0{,}9}{1000}\left(1 - \frac{1000}{10\,000}\right)} = \sqrt{\frac{9}{100\,000} \cdot \frac{9}{10}} =$$

$$= \sqrt{\frac{81}{1000000}} = \sqrt{0{,}000081} = 0{,}009.$$

Tuo būdu skirtumas sudaro $0{,}015 - 0{,}009 = 0{,}006$ arba 0,6%. Tačiau formula XXXVb, kaip paprastesnė, tam tikrais atsitikimais gali būti naudojama. O aplamai, formula XXXV, ją modifikavus, įgalina atsakyti ir į anksčiau pastatytąjį klausimą apie atsitiktinai reikalingąjį reprezentantų arba stebėjimų n skaičių, kad statistinio dalinio tikimybės p' paklaida iš anksto apibrėžtųjų ribų neprašoktų.

Imkime mūsų naudotąją formulą XXXV ir panagrinėkime. Žinome, kad

$$\sqrt{\frac{p\,(1-p)}{n}\left(1 - \frac{n}{N}\right)} = \bar{u}_p,$$

— t. y. duoda skaičių, kuris apibrėžia tikimybės (p) vidutinę (normalinę) paklaidą. Bet mes jau matėme, kad vidutinės arba normalios paklaidos ($\bar{u}$) apibrėžiamos ribos garantuoja tik apie 0,683 arba 68,3% tikimybę, dviejų sigmų (mūsų atveju dviejų $\bar{u}$) ribose tikimybė pasiekia 0,954 arba 95,4% ir trijų sigmų ribose — 0,997 arba 99,7%. Taigi bet kurio st. masės dalinio tikimybinio dydžio (p) griežtumas gali būti apibrėžtas vidutine paklaida ($\bar{u}$). Jei sutarsime reprezentacinių duomenų griežtumą žymėti raide g, o tikimybės laipsnį atitinkantį koeficientą raide t, tai reprezentantų tikslumo laipsnį bus galima apibrėžti pagal šią formulą:

$$g = t\bar{u} = t\sqrt{\frac{p\,(1-p)}{n}\left(1 - \frac{n}{N}\right)}.$$

Bet reprezentacinės statistikos praktikai reikalingas ne tikslumo laipsnis, kuris, kaip jau matėme, kokiu norint griežtumu gaunamas trijų sigmų taisyklės taikymu, o klausimas — kiek gi reikalinga duotomis sąlygomis minimum reprezentantų (n), kad, jais pasirėmus būtų galima tarti, jog st. dalinio tikimybės griežtumas (g) arba tikslumas (t) apibrėžiama vieno ar kito dydžio paklaida? Į šį klausimą atsakymą gali duoti atitinkamai pertvarkyta formula, kur:

$$n = \frac{t^2 Np'(1 - p')}{Ng^2 + t^2 p'(1 - p')} \qquad \text{(XXXVI).}$$

Imkime mūsų jau naudotuosius duomenis, kur dalinio (dį) tikimybė $p' = 0{,}5$, o $N = 10\,000$. Jei reprezentacinio stebėjimo vidutinę paklaidą apibrėšime ū % 0,01, o dalinio p tikimybės tikslumą 0,997, tai tikimybės laipsnį atitinkantį koeficientą sudarys trys sigmos (nes 0,997 tikslumas prie normalios dispersijos gaunamas iš 3ū). Tuo būdu gaunama: $t = 3$, $g = tū = 3 \cdot 0{,}01 = 0{,}03$; formula XXXVI duoda:

$$n = \frac{t^2 Np(1 - p)}{Ng^2 + t^2 p(1 - p)} = \frac{3^2 10\,000 \cdot 0{,}5(1 - 0{,}5)}{10\,000 \cdot (0{,}03)^2 + 3^2 \cdot 0{,}5(1 - 0{,}5)} =$$

$$= \frac{9 \cdot 10\,000 \cdot 0{,}25}{10\,000 \cdot 0{,}0009 + 9 \cdot 0{,}25} = \frac{22\,500}{9 + 2{,}25} = \frac{22\,500}{11{,}25} = 2\,000.$$

Tuo būdu randame, kad reprezentaciniu metodu norint gauti (p) tikimybę 0,997 tikslumo ir vidutinę paklaidą apibrėžti ū = 0,01 arba 1%, — reprezentantų skaičius negali būti mažesnis kaip $n = 2\,000$. Mažesnio reprezentantų skaičiaus n pakaktų tik tuo atveju, jei pasitenkintume kiek didesniu netikslumu t arba mažesniu tikimybės laipsniu. Jei, pav., kitos sąlygos paliktų tos pačios, bet pasitenkintume tik dviejų sigmų aprėžiama tikimybe (0,954), tai $t = 2$ ir tuo būdu gautume n:

$$n = \frac{2^2 \cdot 10\,000 \cdot 0{,}25}{10\,000 \cdot (0{,}03)^2 + 2^2 \cdot 0{,}25} = \frac{4 \cdot 2\,500}{9 + 1} =$$

$$= \frac{10\,000}{10} = 1\,000.$$

Tad šiuo atveju reprezentantų pakaktų 1000 arba 10%; užtat dalinio tikimybė tesiektų tik 95,4%. Jei ieškomos tikimy-

bės artumas tikrenybei paliktų kaip ir anksčiau (0,997), bet pasitenkintume mažesniu griežtumu (didesnės paklaidos nebijotume), sakysim, g = 0,25, tai ta pati formula duotų:

$$n = \frac{3^2 \cdot 10\,000 \cdot 0{,}25}{10\,000 \cdot (0{,}05)^2 + 3^2 \cdot 0{,}25} = \frac{9 \cdot 2\,500}{10\,000 \cdot 0{,}0025 + 9 \cdot 0{,}25} =$$

$$= \frac{22\,500}{25 + 2{,}25} = \frac{22\,500}{27{,}25} (=) 826.$$

Iš čia matome, kad atsisakius griežtumo (tariant, kad didesnė paklaida pavojaus nesudaro), reprezentantų skaičius gali būti dar mažesnis.

Pakeiskime dar dalinio sąlygą, pav., p = 0,1, o griežtumą g ir tikimybę t palikime kaip pirmame uždavinyje, t. y. g = 0,03 ir t = 3. Tada gausime n:

$$n = \frac{3^2 \cdot 10\,000 \cdot 0{,}1 \cdot 0{,}9}{10\,000 \cdot (0{,}03)^2 + 3^2 \cdot (0{,}1 \cdot 0{,}9)} = \frac{9 \cdot 10\,000 \cdot 0{,}09}{10\,000 \cdot 0{,}0009 + 9 \cdot 0{,}09} =$$

$$= \frac{90\,000 \cdot 0{,}09}{9 + 0{,}81} = \frac{8\,100}{9{,}81} (=) 826.$$

Kaip matome, reprezentacinio metodo statistinės masės dalinių aprėžtose ribose griežtumui gauti yra reikalinga juo didesnis reprezentantų skaičius n, kuo dalinys didesnis — ir, žinoma, atvirkščiai.

Dar patikrinkime XXXVI formulos bei paties skaičiavimo tikslumą. Jeigu g = tū, tai

$$\bar{u} = \sqrt{\frac{p\,(1-p)}{n}\left(1 - \frac{n}{N}\right)}.$$

O kadangi žinome, jog g = 0,03 ir t = 3, tai vidutinė paklaida turėtų išeiti $\bar{u} = \frac{0{,}3}{3} = 0{,}01$. Kai iš anksto buvome sutarę, kad p = 0,1 ir N = 10 000, o n (=) 826 davė formula XXXVI, tai pagal formulą XXXVa turėtume gauti:

$$\bar{u} = \sqrt{\frac{p\,(1-p)}{n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{0{,}1 \cdot 0{,}9}{826}\left(1 - \frac{826}{10\,000}\right)} =$$

$$= \sqrt{\frac{0{,}09}{826} \cdot 0{,}9174} = \sqrt{\frac{825{,}66}{8\,260\,000}} (=)$$

$$(=) \sqrt{0{,}0000999} (=) \sqrt{0{,}0001} (=) 0{,}01.$$

c) Reprezentacinių st. vidurkių tikimybė bei paklaidos.

Reprezentacinės statistikos praktikoj gal dažniausia operuojama vidurkiais. Pav., reprezentacinėse statistikos srityse, kur ištisinis stebėjimas beveik nepritaikomas ir kur yra nuo pat pradžios kaip ir koks reprezentacinės statistikos rezervatas arba monopolis, k. a. derlių ir kainų statistikoj, beveik išimtinai operuojama vidurkiais. O vidurkių patikimoji, normali ir galimoji paklaida, kaip jau iš ankstyvesnių skyrių (8, pp. 258—265) parodyta, matuojamos viduriniu kvadratinguoju nukrypimu pagal žinomąją trijų sigmų taisyklę. Vadinasi, ir čia, kaip reprezentacinės statistikos dalinių tikimybei nustatyti, taip pat reikalinga žinoti, kiek turi būti išplėstas reprezentacinis stebėjimas, t. y. kaip didelis turi būti paimtas reprezentantų skaičius n iš tyrinėjamos masės N, kad galimoji vidurkio paklaida prie aprėžtos tikimybės neprašoktų iš anksto apibrėžto dydžio.

Savo esmėje šis klausimas atsakomas tomis pačiomis matematinio analizio priemonėmis, kaip ir klausimas, išaiškintas praeitame skyriuje (b). Skirtumas tik tas, kad šiuo atveju kvadratingajam nukrypimui σ tenka toks pat vaidmuo, kaip XXXV ir XXXVI apibrėžtose formulose reiškiniui $p(1-p)$. Šitai turėdami galvoje, reprezentacinės statistikos vidutinę arba normalią paklaidą galime apibrėžti pagal XXXV analogišką formulą

$$\bar{u} (=) \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)} \qquad \text{(XXXVII).}$$

Jei sutarsime, kaip ir skyriuje b, reprezentacinio vidurkio ribų griežtumą g apibrėžti tais pačiais simboliais pagal formulą

$$g = t\bar{u} = t \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)},$$

tai pagal analogiją su XXXVI formula reprezentantų skaičių n gausime:

$$n = \frac{t^2 N \sigma^2}{Ng^2 + t^2\,\sigma^2} = \frac{t^2 N}{\frac{Ng^2}{\sigma^2} + t^2} \qquad \text{(XXXVIII)}.$$

Šioms formuloms praktikoj pritaikinti paimkime jau naudotuosius duomenis apie kviečių derlius iš 1 ha. Mes žinome, kad bendrasis pasėlių plotas siekė 162 600 ha; tad N = 162 600. Tarkime, kad reprezentaciniu metodu ištyrėme 162,6 ha ir gavome aritmetinį vidurkį $\bar{x} = 12{,}95$ q, o kvadratingąjį nukrypimą $\sigma = 1{,}09$ q. Tuo būdu normalią paklaidą ū pagal XXXVIII formulą gauname:

$$\bar{u} = \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{1{,}09^2}{162{,}6}\left(1 - \frac{162{,}6}{162\,600}\right)} =$$

$$(=) \sqrt{\frac{1{,}19}{162{,}6} \cdot 0{,}999} \,(=)\, \sqrt{\frac{1{,}19}{162{,}6}} \cdot 1 \,(=)\, \frac{1{,}09}{12{,}75} =$$

$$(=) 0{,}08 \text{ (q)}.$$

Jei sutarsime, kad t = 3 (vadinasi vidurkio tikimybė aprėžiama 0,997), tat $g = t\bar{u} = 3 \cdot 0{,}08 = 0{,}24$. Toliau reikalingas reprezentantų skaičius gaunamas pagal formulą XXXVIII:

$$n = \frac{t^2 N \sigma^2}{Ng^2 + t^2\,\sigma^2} = \frac{3^2 \cdot 162\,600 \cdot 1{,}09^2}{162\,600 \cdot 0{,}24^2 + 3^2 \cdot 1{,}09^2} =$$

$$= \frac{9 \cdot 162\,600 \cdot 1{,}19}{162\,600 \cdot 0{,}0576 + 9 \cdot 1{,}19} = \frac{1\,741\,446}{9376{,}47} (=)\ 185{,}7 \text{ ha}.$$

Iš čia matome, kad net pasitenkinant gautosios vidutinės paklaidos ū palyginamuoju didumu ir vidurkio griežtumą aprėžus 2% (kai $g = 3 \cdot 0{,}08 = 0{,}24$, o $\bar{x} = 12{,}95$, tai $\frac{g}{\bar{x}} = \frac{0{,}24}{12{,}95}$ (=) 0,02), iš 162,6 ha stebėjimo duomenų yra permaža. O norint, kad vidurkio vidutinė paklaida neprašoktų 0,01 ir griežtumas (galimoji paklaida) g nebūtų didesnė kaip 3%, iš tos pačios formulos gaunama:

$$n = \frac{3^2 \cdot 162\,600 \cdot 1{,}09^2}{162\,600 \cdot 0{,}03^2 + 3^2 \cdot 1{,}09^2} = \frac{9 \cdot 162\,600 \cdot 1{,}19}{162\,600 \cdot 0{,}0009 + 9 \cdot 1{,}19} =$$

$$= \frac{1\,741\,446}{146{,}34 + 10{,}71} = \frac{1\,741\,446}{157{,}05} (=)\ 11\,088{,}5 \text{ ha}.$$

Jei palikti tą patį griežtumą g = 0,03, o tenkintis 0,954 arba 95,4% tikimybe, tai gausime t = 2, ir n

$$n = \frac{2^2 \cdot 162\,600 \cdot 1{,}09^2}{162\,600 \cdot 0{,}03^2 + 2^2 \cdot 1{,}09^2} = \frac{4 \cdot 162\,600 \cdot 1{,}19}{162\,600 \cdot 0{,}0009 + 4 \cdot 1{,}19} =$$

$$= \frac{773\,976}{146{,}34 + 4{,}76} = \frac{773\,976}{151{,}10} (=)\ 5\,122{,}3 \text{ (ha)}.$$

Žodžiu, reprezentantų skaičiaus n mažinimas privalo būti išpirktas arba didesnės vidutinės bei galimos paklaidos arba mažesnės duomenų tikrumo tikimybės.

d) Orientacinių bei „geriausiųjų" statistinės masės reprezentantų parinkimas.

Iš patiektųjų pavyzdžių matyti, kad reprezentacinių vidurkių, taip pat reprezentacinių st. dalinių tikimybės bei tikslumo ribos ir atsitiktinių reprezentantų minimalinis skaičius n, kuris yra reikalingas apibrėžtos tikimybės bei griežtumo reprezentacinės st. dydžiams gauti, kiekvienu konkretiškai paimtu atsitikimu duodasi nustatomas pagal atitinkamas formulas. Bet tos pačios formulos, statistikos praktikoj naudojamos, yra reikalingos kai kurių skaitmeninių dydžių, kurie turi būti žinomi iš anksto. Pav., kad gauti reprezentacinio st. dalinio vidutinę paklaidą ū pagal vieną ar kitą atitinkamos XXXV formulos varijantą, yra reikalingi skaičiai, kuriais būtų galima pakeisti formulos simbolius N, n ir p. Jei tarsime, kad paruošiamajame skaičiavime n gali būti paimtas bet kuris skaičius savivalingai, — nes galutinis reprezentantų skaičius n galės būti nustatytas empiriškai pagal XXXVI formulą pasikliaujant XXXV formulos duotąja vidutine paklaida ū, — vis dėlto reikalinga bent apytikriai žinoti apie tyrinėjamos st. masės bendrąjį vienetų skaičių N. Be to, turi būti vienu ar kitu skaičiumi apibrėžtas ir st. dalinys p, kurio statistiškai tikimybinį tikslumą (preciziją bei griežtumą) apibrėžti kaip tik yra st. analizio uždavinys.

Ta pati padėtis ir reprezentacinio vidurkio precizijos bei patikimumo apibrėžimuose. Ir čia pagrindinė formula (XXXVII), kad galėtų būti pritaikinta, yra reikalinga skaičių, kurie bent

σ ir N ar n simbolius pakeistų. Bet vidurinis kvadratingasis nukrypimas (σ) tegali būti gautas tik atsižvelgiant st. eilutės dydžių skirtumų. Vadinasi, vienu ar kitu būdu privalo būti sužinotas kuris nors st. masės vienetų skaičius (N ar n) ir vidurinis kvadratingasis nukrypimas anksčiau, negu aprėžiama reprezentacinio vidurkio vidutinė paklaida ū.

Iš pirmo pažvelgimo gali atrodyti, kad tai neišsprendžiama problema. Tikrenybėje šios problemos sprendimas nėra nei taip sunkus, nei nenugalimas. Visų pirma, reprezentacinio metodo panaudojimo klausimas gali kilti dviem atvejais: 1) kai ištisinio stebėjimo metodas dėl vienų ar kitų priežasčių nebepritaikomas (pav., kainų, derlių statistikoj) arba 2) kai ištisinio stebėjimo keliu gautieji duomenys norima reprezentaciniu metodu ištirti orientacijos tikslais (kad statistiškai prasmingesnė ir tikslesnė išeitų st. duomenų klasifikacija).

Pradėkime nuo antrojo atvejo. Jei tyrinėjamoji masė ištisiniu st. metodu kada nors jau buvo stebėta, tai bendrasis tokios st. masės vienetų skaičius N (bent apytikriai) turi būti žinomas. Galimas daiktas, kad tokiais atvejais bus žinomi ir apytikriai st. dalinių bei vidurkių dydžiai. O jeigu ne, tai formuloms reikalingus p, $\bar{x}$, σ ar n dydžius orientacijos reikalui bus lengva gauti pagal vieną ar kitą reprezentacinio metodo sistemą.

Jau sk. 7, b ir c buvo parodyta, jog labai jau maža tikimybė, kad bent kiek didesnį atsitiktinių reprezentantų skaičių paėmus patektų tik kuris nors vienas ar, aplamai, masėje rečiausiai sutinkamas varijantas. O jeigu taip, tai ir formuloms reikalingų o r i e n t a c i n i ų dydžių p, σ ar $\bar{x}$ reiškiniams pakeisti bus lengva gauti. O tais atvejais, kai ištisinis st. metodas iš viso nepritaikomas, tenka vienu ar kitu būdu orientacinių dydžių iš karto ieškoti reprezentaciniu metodu, kuris šiuo atveju atstoja paruošiamąjį arba bandomąjį kontrolinį stebėjimą, paprastai daromą prieš visuotinįjį surašymą (ištisinį stebėjimą). Reiškinys N tada aplamai atkrinta ir reikalinguosius dydžius minimaliniam reprezentantų skaičiui n gauti tokiais atvejais galima išvesti pasikliaujant orientacinio arba preleminarinio stebėjimo gautaisiais st. dalinių tikimybiniais (p) bei vid. kvadratingojo (σ) dydžiais.

Vadinasi, tais atvejais, kai N, σ ar p tikrieji dydžiai net apytikriai nežinomi, patys orientaciniai p ar σ dydžiai neprivalo būti paimti iš pirmojo ir vienintelio stebėjimo, o bent kelių bandomųjų reprezentacinių grupių vidurkis. Iš kitos, tačiau, pusės, visiškas aiškiai išsiskiriančių statistinės masės dalinių bei jų ypatybių paneigimas taip pat gali galutinuosius st. duomenis nuo tikrenybės atitolinti. Be to, operuojant tipiškųjų st. dalinių dydžiais, to paties laipsnio tikimybei bei griežtumui gauti, aplamai, būna reikalingas mažesnis reprezentantų skaičius, negu operuojant atsitiktinių reprezentantų skaičiumi, kaip vienetu.

Paimkime vėl tuos pačius kviečių derliaus duomenis. Psl. 256 atspausdintoj lentelėj 31 yra patiekti 23 apskričių derliaus duomenys. Statistiškai metodologiniu atžvilgiu pavyzdiniam skaičiavimui reprezentacinei statistikai galime sutarti, kad kiekvieno apskrities parodytas 1 ha derlius yra atsitiktinis reprezentantas iš mums tiksliai nežinomo, bet labai didelio, lyginant su n, ha skaičiaus. Vienu ar kitu metodu išskaičiuotas kvadratingasis nukrypimas sulig skirtumais nuo a. vidurkio $\bar{x} = 12{,}98$ duoda σ^2 (=) 2,96, taigi $\sigma = \sqrt{2{,}96}$ (=) 1,72. Pangrinėkime, ar būtų kokia prasmė nesitenkinti σ, o ieškoti $\bar{\sigma}$, t. y. iš tų pačių 23 reprezentantų sudarytųjų grupių vidurkio. Pasielgti galime dvejopai: 1) sudaryti, sakysime, 4 atsitiktines grupes burtų keliu, imdami po šešis derliaus variantus į kiekvieną, arba 2) sudaryti tokį pat „tipiškųjų" grupių skaičių imdami pagrindan: a) vienodą derliaus intervalą ir b) sudarydami grupes vad. Galton'o metodu.

Kadangi $n = 23$, o grupių norime turėti keturias, tai vienai grupei tektų tik penki variantai.

Sutarkime pirmajam variantui reikalingas atsitiktines grupes sudaryti pagal abėcėlę. Surašę kiekvienos apskrities derlius abėcėlės tvarkoj pradedant Alytaus ir baigiant Zarasais gauname šitokias grupes ir jų vidurkius:

$$\text{I.}\quad \bar{x}_1 = \frac{11{,}9 + 13{,}6 + 11{,}6 + 13{,}3 + 11{,}9 + 14{,}7}{6} \text{ (=) } 12{,}83;$$

$$\text{II.}\quad \bar{x}_2 = \frac{13{,}0 + 13{,}6 + 12{,}7 + 18{,}2 + 13{,}4 + 13{,}3}{6} \text{ (=) } 14{,}03;$$

III. $\bar{x}_3 = \frac{11{,}2 + 12{,}6 + 13{,}7 + 15{,}7 + 13{,}3 + 10{,}5}{6}$ (=) 12,83;

IV. $\bar{x}_4 = \frac{11{,}3 + 11{,}5 + 11{,}0 + 15{,}1 + 11{,}4}{5}$ (=) 12,06.

Tuo būdu gauname keturis atsitiktinius reprezentantų vidurkius $\bar{x}$; ir iš jų bendrąjį arba generalinį (vyriausią) reprezentantų vidurkį:

$$\bar{x}'_a = \frac{12{,}83 + 14{,}03 + 12{,}83 + 12{,}06}{4} = 12{,}94.$$

Atsižvelgiant kiekvienos grupės lyginamojo svorio — vidurkis $\bar{x}$ (=) 12,98.

Atsitiktinių reprezentantų grupių kvadratingieji nukrypimai gaunami pagal

$$\sigma_1 = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2} =$$

$$= \sqrt{\frac{11{,}9^2 + 13{,}6^2 + 11{,}6^2 + 13{,}3^2 + 11{,}9^2 + 14{,}7^2}{6} - 12{,}83^2} (=)$$

$$= \sqrt{\frac{995{,}72}{6} - 12{,}83^2} (=) \sqrt{165{,}95 - 164{,}61} = \sqrt{1{,}34} (=) 1{,}16;$$

Panašiu būdu gausime:

$$\sigma_2 = \sqrt{\frac{1202{,}94}{n} - 14{,}03^2} (=) \sqrt{200{,}49 - 196{,}84} = \sqrt{3{,}65} (=) 1{,}91;$$

$$\sigma_3 = \sqrt{\frac{1005{,}62}{6} - 12{,}83^2} = \sqrt{167{,}59 - 164{,}61} = \sqrt{2{,}98} (=) 1{,}73;$$

$$\sigma_4 = \sqrt{\frac{738{,}91}{5} - 12{,}06^2} = \sqrt{447{,}78 - 145{,}44} = \sqrt{2{,}34} (=) 1{,}53.$$

Tuo būdu XXXVIIa formulai reikalingąjį reprezentacinių grupių vidutinį kvadratingąjį nukrypimą σ gauname:

$$\bar{\sigma} = \frac{1{,}16 + 1{,}91 + 1{,}73 + 1{,}53}{4} (=) 1{,}57 \text{ q.},$$

Jei gi atsižvelgti kiekvienos σ lyginamojo svorio, tai gausime:

$$\bar{\sigma} = \frac{1{,}16 \cdot 6 + 1{,}19 \cdot 6 + 1{,}73 \cdot 6 + 1{,}53 \cdot 5}{6 + 6 + 6 + 5} = \frac{36{,}45}{23} (=) 1{,}58.$$

„Tipiškųjų" reprezentantų grupių sistemoj pirmasis variantas (statistinės grupės vienodais intervalais) bus reikalingas intervalo iš 2 ha, nes $\frac{18 - 10}{4} = 2$. Tuo būdu gausime:

I. 10 — 12 ha:

$$\bar{x}_1 = \frac{10{,}5 + 11{,}0 + 11{,}2 + 11{,}3 + 11{,}4 + 11{,}5 + 11{,}6 + 2 \cdot 11{,}9}{9} = (=) 11{,}37$$

II. 12 — 14 ha:

$$\bar{x}_2 = \frac{12{,}6 + 12{,}7 + 13{,}0 + 3 \cdot 13{,}3 + 13{,}4 + 2 \cdot 13{,}6 + 13{,}7}{10} = 13{,}25;$$

III. 14 — 16 ha: $\bar{x}_3 = \frac{14{,}7 + 15{,}1 + 15{,}7}{3} (=) 15{,}17;$

IV. 16 — 18 ha: $\bar{x}_4 = \frac{18{,}2 + 0}{1 + 0} = 18{,}2.$

$$\bar{x}' = \frac{11{,}37 \cdot 9 + 13{,}25 \cdot 10 + 15{,}17 \cdot 3 + 18{,}2}{23} = \frac{298{,}53}{23} (=) 12{,}98.$$

Vidurinis kvadratingasis nukrypimas išeina:

$$\sigma_1 = \sqrt{\frac{x^2}{n_1} - \bar{x}_1^2} =$$

$$= \sqrt{\frac{1164{,}37}{9} - 11{,}37^2} (=) \sqrt{129{,}37 - 129{,}28} (=) \sqrt{0{,}09} = 0{,}30;$$

$$\sigma_2 = \sqrt{\frac{1756{,}65}{10} - 13{,}25^2} (=) \sqrt{175{,}67 - 175{,}46} = \sqrt{0{,}21} (=) 0{,}46;$$

$$\sigma_3 = \sqrt{\frac{690{,}59}{3} - 15{,}17^2} (=) \sqrt{230{,}20 - 230{,}13} = \sqrt{0{,}07} (=) 0{,}26.$$

σ_4, kai $x_{14} = \bar{x}_{14}$ (tėra tik vienas variantas iš viso), bus lygus 0.

Tuo būdu tipiškųjų statistinių grupių sulig 1 ha derlingumu gautųjų kv. nukrypimų σ_i vidurkis $\bar{\sigma}$ išeina:

$$\bar{\sigma}a = \frac{0{,}30 \cdot 9 + 0{,}46 \cdot 10 + 0{,}26 \cdot 3 + 0 \cdot 1}{9+10+3+1} =$$

$$= \frac{3{,}60 + 4{,}60 + 0{,}78 + 0}{23} (=) 0{,}35.$$

Jastremskio ir Chotinskio vadovaujamas Maskvos jaunųjų statistikų kolektyvas reprezentacinių grupių vidurinį kvadratingąjį nukrypimą skaičiuoja betarpiškai iš kiekvienos tipiškos grupės nukrypimų kvadratų bendrosios sumos pagal formulą

$$\bar{\sigma}^2 = \frac{1}{n} \Sigma \sigma_i^2.$$

Mūsų pavyzdyje, kai $\sigma_1^2 = 0{,}09$, $\sigma_2^2 = 0{,}21$, $\sigma_3^2 = 0{,}07$, $\sigma_4^2 = 0$ ir $n = 4$, gauname:

$$\bar{\sigma}^2 = \frac{0{,}09 + 0{,}21 + 0{,}07 + 0{,}00}{4} = \frac{0{,}38}{4} (=) 0{,}09.$$

O kai $\bar{\sigma}^2 = 0{,}09$, tai $\bar{\sigma} = \sqrt{0{,}09} = 0{,}3$, t. y. dar mažesnis. Tačiau, šios rūšies vidutinė sigma $\bar{\sigma}$ yra dirbtinis dydis, kuriuo negalima remtis nustatant tikimybės laipsnį nei paklaidos ribas v i s ų vad. tipiškųjų grupių reprezentuojamai masei (jei tipiškųjų grupių intervalą pakankamai aprėžti, grupinių $\bar{x}$ ir x_i skirtumai išnyktų, o tuo pačiu ir σ prisiartintų 0).

Dėl to yra jau prasmingiau $\bar{\sigma}$ skaičiuoti sulig skirtumais nuo tipiškųjų grupių vidurkių $\bar{x}_i$ pagal formulą

$$\bar{\sigma} = \sqrt{\frac{\bar{x}_i^2}{n} - \bar{x}'^2}.$$

Šios rūšies vidurinis kvadratingasis nukrypimas gaunamas $\bar{\sigma} (=) 1{,}69$, t. y. mažesnis kaip iš reprezentantų visumos, kur $\sigma (=) 1{,}72$, bet didesnis kaip iš tų pačių 23 reprezentantų atsitiktinių grupių, kur, kaip matėme, $\bar{\sigma} (=) 1{,}58$.

Galton'o metodu sudarytųjų dydžio grupių vidurkiai ir jų nukrypimai savo esme nesiskiria nuo tipiškųjų grupių. Vadinasi, ir čia statistiškai prasmingas $\bar{\sigma}$ gali būti sudaromas tik sulig skirtumais iš grupinių vidurkių. Jį išskaičiavus, gaunama $\bar{\sigma} (=) 1{,}62$.

Gautieji duomenys rodo, kad orientaciniai reprezentacinės statistikos dydžiai (šiuo atveju vid. kvadratingieji nukrypimai) iš to paties st. stebėjimo ir tų pačių reprezentantų žymiai skiriasi sulig tuo, ar jie gaunami iš reprezentantų visumos betarpiškai, ar suskirsčius reprezentantus į tipiškas st. grupes, ar pagaliau, iš atsitiktinių tokio pat skaičiaus grupių. Reprezentacinės statistikos praktikoje tai yra svarbi aplinkybė.

Patikrinkime. Pusl. 280 yra parodyta, kad iš 162,6 ha kviečių derliaus stebėjimų, kai bendrasis pasėlių plotas siekia 162 600 ha, gaunama vidutinė vidurkio paklaida $\bar{u} = 0{,}08$. Norint šią paklaidą sumažinti ligi $\bar{u} = 0{,}01$ ir gauti bendrojo derliaus tikimybę su griežtumu 99,7% (galimąja paklaida nedidesne kaip 0,03), atsitiktinių reprezentantų skaičių n tenka padidinti ligi 11 089 ha (psl. 280). Jeigu remtis tik 23 reprezentantų stebėjimu ($n = 23$), tai $\sigma = 1{,}72$ ir vidutinė paklaida:

$$\bar{u} = \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{1{,}72^2}{23}\left(1 - \frac{23}{162\,600}\right)} (=)$$

$$(=) \sqrt{\frac{2{,}96}{23} \cdot 1} (=) \sqrt{0{,}13} (=) 0{,}36.$$

Šiomis sąlygomis 99,7% tikimybės, kai $t = 3$, griežtumas (galimoji paklaida) $g = tu = 3 \cdot 0{,}36 = 1{,}08$ q. O tai jau sudaro apie 8,3% vidurkio galimą paklaidą $\left(\frac{1{,}08}{12{,}98} (=) 0{,}083\right)$. Norėdami vidurkio galimą paklaidą aprėžti $g = 0{,}03$, o vidutinę $\bar{u} = 0{,}01$, reikalingąjį atsitiktinių reprezentantų skaičių pagal formulą XXXVIII gausime:

$$n = \frac{t^2 N \sigma^2}{Ng^2 + t\sigma^2} = \frac{3^2 \cdot 162\,600 \cdot 1{,}72^2}{162\,600 \cdot 0{,}03^2 + 3^2 \cdot 1{,}72^2} =$$

$$= \frac{9 \cdot 162\,600 \cdot 2{,}96}{162\,600 \cdot 0{,}0009 + 9 \cdot 2{,}96} = \frac{1\,463\,400 \cdot 2{,}96}{146{,}34 + 26{,}64} =$$

$$= \frac{4\,331\,664}{172{,}98} (=) 25\,041 \text{ ha iš } 162\,600 \text{ ha bendrojo pasėlių ploto.}$$

Jei skaičiavimo pagrindan dėsime iš tų pačių 23 reprezentantų sudarytųjų atsitiktinių grupių sigmos vidurkį, kuris, kaip

žinome, siekia $\bar{\sigma} = 1{,}58$, tai XXXVIII formula to paties laipsnio bei griežtumo tiekybei reikalingąjį n parodys mažesnį kaip 25 000, bet vis dėlto reprezentantų skaičius n dar turės būti žymiai didesnis kaip 11 000, kuris yra gautas pasikliaujant $\sigma = 1{,}09$, t. y. visu trečdaliu mažesniu kvadratinguoju nukrypimu.

Kai dėl orientacinių st. dydžių reprezentacinei statistikai gavimo ir tam tikslui orientacinių reprezentantų bei orientacinių st. masės tipų sudarymo, tai jų reikšmė yra nemažiau svarbi. Nes sulig jais gautieji dydžiai bei tikimybiniai skaičiai, o pirmoj eilėj nustatytasis pagal XXXVIII formulą minimalinis reprezentantų skaičius n yra tik tiek patikimas, kiek orientacinių st. grupių tipai savo turiniu ir lyginamuoju svoriu atitinka visos st. masės tipiškųjų dalinių turinį bei lyginamąjį svorį. Vadinasi, jeigu mes įvairiais statistinio analizio metodais konstatavome, kad orientaciniai dydžiai iš 23 atsitiktinių reprezentantų, suskaidžius juos į 4 tipiškas grupes, duoda tokį kvadratingąjį nukrypimą, kuris garantuoja 0,997 tikimybę prie 0,03 galimos paklaidos tik iš 2—3 000 reprezentacinio (h) stebėjimo, o 162 ar 163 atsitiktinių reprezentantų kvadratingasis nukrypimas tai pačiai tikimybei bei griežtumui rodo esant reikalinga reprezentacinių ha 20—30 000, tai iš to negalima daryti išvados, kad, aplamai, išplėsti stebėjimą nereikalinga. Atvirkščiai, pagal tipiškųjų grupių gautąjį vid. kvadratingąjį nukrypimą orientaciniame skaičiavime apibrėžtas minimalinis reprezentantų skaičius galioja tik tąja sąlyga, jei ir galutinis reprezentacinis stebėjimas taip pat bus tais pačiais principais (tos pačios rūšies tipiškomis grupėmis) suorganizuotas, ir jei, be to, visoje tiriamoje masėje tais pačiais pažymiais sudarytų tipiškųjų grupių lyginamasis svoris atitiks orientacinių grupių lyginamąjį svorį. Vadinasi, didesnė paklaida, negu orientaciniam arba paruošiamajame skaičiavime, vis dar yra galima. Dėl to patsai orientacinių tipų bei juos atitinkančių absoliučių bei vidutinių dydžių sudarymas yra reikalingas atsargaus pasiruošimo. Geriausia tokiais atvejais nesitenkinti vieno ir vienintelio orientacinio stebėjimo duomenimis, o stebėjimą pakartoti keletą kartų. Taip pat ir vad. reprezentacinius tipus rekomenduojama sudaryti, jei tik galima, keliais skir-

tingais kriterijais ir iš skirtingo atsitiktinių reprezentantų skaičiaus. Tuo būdu galima prieiti prie „geriausiųjų“ reprezentantų ir prie patikimiausių skaičių.

Aplamai, nereikia užmiršti, kad tiek orientacinės, tiek galutinės tipiškosios grupės reprezentaciniame stebėjime yra reikalingos kiek galint arčiau tikrenybės esančių svorio koeficientų, o ne mechaniškai sudarytų grupių iš vienodo reprezentantų skaičiaus. Kai taip yra, tai ir anksčiau patiektosios vidutinės paklaidos ir kt. reprezentacinio metodo formulos yra reikalingos pataisų. Šios pataisos savo esmėje yra visai analogiškos pataisoms, kurios reikalingos svertiniams vidurkiams, nevienodai pasiskaidžiusios masės kvadratingiesiems nukrypimams ir kt. (žiūr. formulas VIIa, XXIIb ir kt.). Vidutinei paklaidai ir kt. dydžiams gauti galioja šitoki variantai:

$$\sigma = \sqrt{\frac{\Sigma (x_j^2 y_j)}{\Sigma y_j} - \bar{x}_j^2} = \frac{\Sigma (\Delta_j^2 y_j)}{\Sigma y_j}$$

$$\bar{\sigma} = \frac{\Sigma \sigma_j y_j}{\Sigma y_j} = \frac{\Sigma \sigma_j y_j}{n}$$

$$\bar{u} = \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)} (=) \sqrt{\frac{\sigma^2}{n}}$$

Šiomis sąlygomis vidutinės paklaidos II ir XXXV formula gauna naują variantą, kur:

$$\bar{u}_p = \sqrt{\frac{pq}{n}} = \sqrt{\frac{p\,(1-p)}{n}} \qquad \text{(XXXVc).}$$

Atitinkantis XXXVII formulos variantas rašomas:

$$\bar{u}_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}} \qquad \text{(XXXVIIa).}$$

Lengva įsitikinti, kad praktiškai ir kiti XXXV bei XXXVII formulų variantai beveik nesiskiria nuo XXXVc ir XXXVIIa visais tais atvejais, kai n santykyje su N yra labai mažas skaičius. Pav., psl. 279 patiektasis skaičiavimo pavyzdys pagal formulą XXXVII

$$\bar{u} = \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)}$$

praktiškai davė ū pagal formulą XXXVII a, nes išėjo:

$$\bar{u}\,(=)\,\sqrt{\frac{\sigma^2}{n}\,0{,}999\ldots}\,(=)\,\sqrt{\frac{\sigma^2}{n}\,.\,1} = \sqrt{\frac{\sigma^2}{n}}.$$

Vadinasi, ū mažai tepasikeistų, jei turėdami n, savavalingai paimtume labai didelį, lyginant su n, skaičių N, arba žinodami N, paimtume mažytį skaičių n.

Įsižiūrėję į XXXVc ir XXXVIIa formulų variantus mes matome, kad iš kitų skyrių jau gerai pažįstamieji reiškiniai $p\,(1-p)$ ir σ^2 viršuje atžymėti brūkšneliais $\overline{pq} = \overline{p\,(1-p)}$ ir $\overline{\sigma^2}$. O mūsų simbolių žymėjimo sistemoje viršuje uždėti brūkšneliai, paprastai, rodo, kad simbolis atstovauja vidurkį.

Literatūra ir papildomos žinios su skaičiavimo pavyzdžiais ir t. t. galima rasti visuose matematinės statistikos vadovėliuose. Tarp kita ko galima rekomenduoti: 1) H. P. Rietz rašinius „Stichprobenerhebungen" ir „Bernoullische, Poissonsche u. Lexissche Verteilungen" (Rietz-Bauer, Handbuch der Mathematischen Statistik, pp. 93—122), 2) G. U. Yule, An Introduction to the Theorie of Statistics, Part III (Theorie of sampling, pp. 254—356), 3) O. N. Anderson, Einführung in die mathematische Statistik (ypatingai pp. 128—266), 4) Bojarskij ir kt., Statistika (ypač V ir VI skyriai, pp. 124—205), 5) March, L., Les Principes de la Methode Statistique, pp. 557—574 ir kt.

Be to, reprezentacinės statistikos teorijos ir metodų klausimai gana dažnai nagrinėjami statistikos mokslo žurnaluose. Ypač dėmesio vertas Tor Jerneman, On the Substitution of Constans in Sampling Formulae with special Reference to the Theorie of the „too small values" (Nordic Statistical Journal, Vol. 3, pp. 85—112, Stockholm 1931).

10. Interpoliacija.

a) Interpoliacijos esmė.

Istoriškai aprašomosios statistikos vadovėliuose statistinės interpoliacijos metodai neliečiami ar neigiamai apibūdinami (G. Mayr, Mueller) arba susiejami su statistinių eilučių išlyginimu bei pagrindinės linkmės nustatymu (Žižek ir kt.). Tačiau interpoliacijas sieti su išlyginimais formaliai tiek tėra pagrindo, kad interpoliacijoms kai kada naudojamos tos pačios tiesės bei

parabolinės lygtys, taip pat aritmetinis, geometrinis, diferencinis, sumavimų ir kt. metodai, kurie ir statistikos eilučių ar linkmės nustatymams naudojami. O interpoliacijų statistinis uždavinys savaime yra visai kitas, ne kaip statistinio eilučių išlyginimo.

Jau žvilgsnis į abiejų statistinio analizio metodų techniką turi parodyti jų esminį skirtumą. Kai statistinis išlyginimas ar pagrindinės linkmės nustatymas paprastai operuoja s t a t i s t i n ė m i s, t. y. d a u g e l i o skaičių eilutėmis, tai interpoliuojant dažniausiai turima reikalo su d v i e m pagrindiniais skaičiais, ir tik, palyginti, retais atsitikimais socialekonominėj statistikoj interpoliuojama naudojant daugiau kaip du pagrindinius skaičius. Ši aplinkybė kai kam duoda pagrindo net abejoti dėl interpoliacijos statistikai prasmingumo (statistika turi reikalo su m a s ė m i s arba m a s i n i a i s reiškiniais ir esą statistinės prasmės tegali turėti tik s t a t i s t i n i ų, t. y. didžiųjų skaičių dėsnio įtakoje esančių, o ne paprastųjų aritmetinių eilučių analizis).

Vis dėlto statistinio analizio darbas būtų visai nereikalingai susiaurintas ir pati statistika nualinta, jei interpoliacijos teikiamųjų metodų pagalbos statistikoj būtų atsisakyta. Visų pirma, jau ieškant mediano lyginių (porinių) skaičių eilutėj interpoliuojama: nes kai statistinės eilutės dydžių skaičius n yra lyginis, tai M (medianas) gaunamas iš abiejų vidurinių dydžių aritmetinio vidurkio, t. y. lygiai taip pat, kaip gaunamas patsai vidurinis dydis vad. vidurkinėj interpoliacijoj. Šiuo atveju medianas tik tuo atveju atitiks statistinės eilutės konkretiškai paimtą dydį, kai abu viduriniai dydžiai bus lygiaverčiai[9]. Taip pat tik interpoliacijos keliu galima nustatyti medianą arba kitą kurį nors statistinį dydį (kvartilį, decilį ar kt.) ir tais atvejais, kai statistinė eilutė yra suskirstyta dydžio grupėmis su mažesniais ar didesniais intervalais.

Statistinės interpoliacijos esmė ir prasmė bus lengviausia suprasti prisiminus interpoliacijos metodus iš matematikos ele-

---

[9] Jei sutarsime lyginių skaičių eilutės vidurio gretimuosius dydžius žymėti x_m ir x_n tai:

$$M_x = x_m + \frac{x_n - x_m}{2} = \frac{x_n + x_m}{2}.$$

mentoriaus. Skaičiavimams palengvinti naudojant logaritmų tabeles, dažnai tenka interpoliacijos keliu rasti logaritmai arba logaritmus atitinkantieji skaičiai visais tais atvejais, kai ieškomųjų dydžių tabelėj nėra, bet yra griežtai aprėžtame santykyje du artimieji dydžiai. Pav., jei mums reikalinga surasti skaičių, atitinkantį logaritmą 1,3112, tai mokykloj naudojamoj dešimtainių logaritmų tabelėje betarpiškai tokio skaičiaus nerasime. Apibrėžtajam artimiausi logaritmai yra 1,3010 ir 1,3222, iš kurių pirmasis atstovauja skaičių n = 20, o antrasis n = 21. Abiejų logaritmų skirtumas čia siekia 0,0212 (1,3222—1,3010 = = 0,0212). Visai pagrįstai darome išvadą, kad šis skirtumas atitinka gretimųjų skaičių 20 ir 21 skirtumą, t. y. 1. Tuo tarpu mūsų logaritmas turi skaitmenis 1,3112, t. y. už n = 20 logaritmą yra didesnis 0,0108. Jei skirtumas tarp 20 ir 21, t. y. 1, yra sudaręs logaritmų mantisės skaitmeninį skirtumą iš 212, tai šių skaičių pareinamybė gali būti pareiškiama trupmenoj iš $\frac{108}{212}$. Vadinasi ir mums rūpimas skaičius $n = 20 + \frac{108}{212}$, t. y. 20,509.

b) Interpoliacijos metodai.

1) Grafiškoji bei aritmetinė interpoliacija tiesės lygtimi.

Paprasčiausia ir lengviausia yra vad. tiesinė interpoliacija, kuri gali būti atliekama ir grafiškoj koordinatų sistemoj. Du žinomus dydžius įstačius aritmetinio mastelio koordinatų sistemon, t. y. jiems atitinkančias vietas atžymėjus taškais, nubrėžtoji tiesė rodys vietas, atitinkančias bet kuriam intervalui interpoliuotiną dydį.

Skaičiuojant interpoliuotinus dydžius tikslingiausia yra pasikliauti tiesės lygtimi

$$y = a + bx,$$

kur a gali būti pirmasis žinomas dydis, t. y. $a = y_0$, ir b konstanta gaunama iš skirtumo tarp pirmojo ir antrojo žinomo dydžio, padalinto interpoliuotinų intervalų skaičiumi n, pagal formulą:

$$b = \frac{y_n - y_0}{n}.$$

Pavyzdiniam skaičiavimui galime paimti Lietuvos gyventojų skaičius 1923 ir 1937 m. galui iš CSB Statistikos Metraščio. Apvaliais skaičiais 1923 m. Lietuvoj gyventojų buvo 2 175 000, o 1937 m. — 2 550 000. Tad gauname:

$$y_0 = 2\,175\,000, \quad y_n = 2\,550\,000 \text{ ir}$$

$$n = 1937 - 1923 = 14.$$

Tuo būdu b pagal formulą:

$$b = \frac{y_n - y_0}{n} = \frac{2\,550\,000 - 2\,175\,000}{14} = \frac{375\,000}{14} (=)$$

$$(=) 26\,786.$$

Pakeitę tiesinės lygties simbolius skaičiais gauname:

$$y = a + bx = 2\,175\,000 + 26\,786x.$$

Lentelė 32.

LIETUVOS GYVENTOJŲ SKAIČIAUS TIESINĖ (ARITMETINĖ) INTERPOLIACIJA 1924—1927 METAMS (1000 gyv.): y=a+bx=2 175+26,786x.

	Metai		Tikrieji gyv. skaičiai	Interpoluotieji gyventojų skaičiai	Skirtumas	Skirtumų koeficientai % %
x	x_i	x	y	y′=y+bx=2175+ +26,786x	y−y′	$\frac{(y-y') \cdot 100}{y}$
0	0	1	2	3	4	5
1	1923	0	2 175	2 175	0	0
2	1924	1	2 202	2 203	—1	—0,05
3	1925	2	2 229	2 230	—1	—0,04
4	1926	3	2 255	2 259	—4	—0,18
5	1927	4	2 282	2 286	—4	—0,17
6	1928	5	2 309	2 317	—8	—0,35
7	1929	6	2 336	2 340	—4	—0,17
8	1930	7	2 363	2 367	—4	—0,17
9	1931	8	2 389	2 393	—4	—0,17
10	1932	9	2 416	2 422	—6	—0,25
11	1933	10	2 443	2 451	—8	—0,33
12	1934	11	2 470	2 476	—6	—0,24
13	1935	12	2 496	2 500	—4	—0,16
14	1936	13	2 523	2 527	—4	—0,16
15	1937	14	2 550	2 550	0	0

Iš patiektosios lentelės matyti, kad aritmetinė gyventojų skaičiaus interpoliacija nė vieniems metams per visą laikotarpį nedavė paklaidos, kuri būtų pasiekusi 0,5% (didžiausia paklaida tenka 1933 m., kai skirtumas sudarė 0,3%). Kad interpoliuotasis y'_{14} atitinka y_{14}, tai savaime suprantamas dalykas.

Jeigu atsitiktų, kad $y_n < y_0$ (paskutinis skaičius mažesnis už pirmąjį), tai konstanta b išeitų su neigiamu reiškiniu ir ta pati formula pavirstų į $y = a - bx$. Kaip matome aritmetinė (taip pat grafiškoji) tiesės lygtimi interpoliacija yra pagrįsta dėsniu, kad koordinatų sistemoj kiekvienas aritmetinio mastelio tarpelis, arba lygčių sistemoj kiekvienas x dydis, kinta vienu ir tuo pačiu vienetų skaičiumi.

2) Geometrinė interpoliacija pagal sudėtinių procentų formulą.

Aritmetinė interpoliacija nėra tinkama tais atvejais, kai yra pagrindo tarti, jog dviejų turimų dydžių tarpelio linkmės patikimiausias kitimas vyko ne aritmetinėj, o geometrinėj proporcijoj, t. y. vienodu koeficientu. Pav., procentų procentams padėtojo kapitalo suma, iš anksto sutarus aprėžtam laikui nekintamą palūkanų procentą, gali būti surandama su cento griežtumu bet kuriam metui pagal žinomą sudėtinių procentų formulą, kurią mūsų naudojamų simbolių sistemoj galime šitaip rašyti:

$$y_i = y_0 \cdot (1 + p)^i. \qquad \text{(XXXIX).}$$

Šiai formulai reikalinga koeficiento konstanta $1 + p$ yra iš sudėtinių procentų lygčių sistemos pažįstamas reiškinys $1 + p$ (arba $1 + i$), kur p (arba i) paprastai simbolizuoja palūkanų procentą. Tais atvejais, kai interpoliuotinoj linkmėj kitimų koeficientas nežinomas, o yra žinomi tik patys kraštutiniai dydžiai y_0 ir y_n, reikalingam koeficientui nustatyti galioja ir iš sudėtinių procentų skaičiuotės žinomoji formula:

$$1 + p = \sqrt[n]{\frac{y_n}{y_0}}, \text{ arba}$$

$$\lg (1 + p) = \frac{1}{n} (\lg y_n - \lg y_0). \qquad \text{(XXXIXa).}$$

Artutiniais skaičiais interpoliuojant geometriniu metodu galima, naudojant logaritmus, interpoliacijos formulą išreikšti tiesės lygtimi, nes:

$$\lg y = \lg y_0 + \frac{1}{n}(\lg y_n - \lg y_0)x, \text{ arba } Y = A + BX,$$

$$\text{kur } Y = \lg y; \quad A = \lg y_0; \quad B = \frac{1}{n}(\lg y_n - \lg y_0);$$

$X = x$; X ir Y yra čia tiesės bet kurio taško koordinatos.

Pavyzdiniam skaičiavimui sutarkime, kad $y_0 = 100$, $y_n = y_{10} = 120$. Tuo būdu:

$$y = y_0 (1 + p)^n = 100 (1 + p)^n, \text{ o}$$

$$1 + p = \sqrt[n]{\frac{y_n}{y_0}} = \sqrt[10]{\frac{120}{100}} = \sqrt[10]{1{,}2},$$

$$\lg (1 + p) = \frac{1}{n} (\lg y_n - \lg y_0) = \frac{1}{10}(\lg 120 - \lg 100) =$$

$$= \frac{1}{10} (2{,}0792 - 2{,}0000) = \frac{0{,}0792}{10} = 0{,}00792,$$

o atlogaritmavus $1 + p$ (=) 1,0 184.

Tuo būdu gauname lygtis:

$$y = y_0 (1 + p)^n (=) 100 (1{,}0184)^n \text{ arba}$$

$$\lg y = \lg 100 + \frac{1}{10} (\lg 120 - \lg 100) \cdot x.$$

Simbolis x atstovauja y-inės skaičių eilutės n. Tuo tarpu tiems patiems skaičiams interpoliacija tiesės lygtimi duoda:

$$y = a + bx = 100 + \frac{120 - 100}{10} x = 100 + 2x.$$

Pavyzdinio skaičiavimo rezultatai lentelėje šitaip atrodys:

Lentelė 33.

ARITMETINĖS IR GEOMETRINĖS INTERPOLIACIJOS PAVYZDYS:

$y_0 = 100$, $y_n = 120$ ir $n = 10$.

x	$y = a + bx = y_0 + bx$	$y = (1 + p)^n . y_0$
	$100 + 2x = y'$	$1{,}0184^n . 100 = y'$
y_0	$= 100 + 2.0 = 100$	$1{,}0184^0 . 100 (=) 100{,}0$
y_1	$= 100 + 2.1 = 102$	$1{,}0184^1 . 100 (=) 101{,}8$
y_2	$= 100 + 2.2 = 104$	$1{,}0184^2 . 100 (=) 103{,}7$
y_3	$= 100 + 2.3 = 106$	$1{,}0184^3 . 100 (=) 105{,}6$
y_4	$= ..... = 108$	$.....(=) 107{,}6$
y_5	$= ..... = 110$	$.....(=) 109{,}5$
y_6	$= ..... = 112$	$.....(=) 111{,}6$
y_7	$= ..... = 114$	$.....(=) 113{,}6$
y_8	$= ..... = 116$	$.....(=) 115{,}7$
y_9	$= ..... = 118$	$.....(=) 117{,}8$
y_{10}	$= 100 + 2.10 = 120$	$1{,}0184^{10} . 100 (=) 120{,}0$

Lentelė rodo, kad geometrinė interpoliacija pagal sudėtinių procentų formulą y′-inei eilutei duoda skaičius, kurie interpoliuojamo periodo pradžioje atsilieka nuo y′-inės eilutės skaičių, interpoliuotų vad. aritmetiniu metodu. Bet juo toliau, tuo abiem metodais gaunamieji skaičiai darosi vis artimesni, ir, žinoma, pasiekus y_n visai sutampa (y'_n gautasis skirtumas, lyginant su y_n, pareina nuo apytikrinio skaičiavimo).

Sekančioj lentelėj pavaizduota geometrinė interpoliacija pasikliaujant logaritmų.

Palyginus 33 ir 34 lentelių duomenis, matyti, kad geometrinio metodo abudu variantai duoda tuos pačius rezultatus (gautuose skaičiuose maži skirtumai pareina nuo artutinio skaičiavimo technikos apibrėžiant trupmenas).

Vadinasi, kurį geometrinės interpoliacijos variantą naudoti — tai technikinių skaičiavimo patogumų klausimas. Savaime suprantama, kad jeigu $y_n < y_0$ (y_n yra mažesnis už y_0),

Lentelė 34.

GEOMETRINĖS INTERPOLIACIJOS PAVYZDYS LOGARITMŲ PAGALBA.

Sąlygos:

$y_0 = 100$, $y_n = 120$ ir $n = 10$.

$$\lg y = \lg y_0 + \frac{1}{n}(\lg y_n - \lg y_0)x = \lg y' = \lg 100 + \frac{1}{10}(\lg 120 - \lg 100)x$$

x	$2{,}0000 + 0{,}00792\,x = \lg y'$	(y')
0	$2{,}0000 + 0{,}00792 \cdot 0 = 2{,}0000$	100
1	$2{,}0000 + 0{,}00792 \cdot 1 = 2{,}00792$	101,8
2	$2{,}0000 + 0{,}00792 \cdot 2 = 2{,}01584$	103,7
3	$2{,}0000 + 0{,}00792 \cdot 3 = 2{,}02376$	105,6
4	..... $= 2{,}03168$	107,6
5	..... $= 2{,}03960$	109,5
6	..... $= 2{,}04752$	111,6
7	..... $= 2{,}05544$	113,6
8	..... $= 2{,}06336$	115,7
9	..... $= 2{,}07128$	117,8
10	$2{,}0000 + 0{,}00792 \cdot 10 = 2{,}0792$	120,0

tai interpoliuojamieji skaičiai ir pati kitimų linkmė gaunama atvirkščioj proporcijoj. Koordinatų sistemoj vaizduojant grafiškai, aritmetinė interpoliacija visais atvejais duoda t i e s ę, o geometrinės interpoliacijos skaičiais nubrėžtoji linija, pradedant y_1, eina žemiau ir, pamažu kildama, pačiam gale prisijungia prie y_n kartu su tiesia (jeigu $y_0 > y_n$, geometrinės interpoliacijos metodu nubrėžtoji linija atitinkamoj proporcijoj eina aukščiau). Ši aplinkybė pirštu prikišamai rodo, kad geometrinė interpoliacija gali būti prasminga tik tais atvejais, kai interpoliuotinų dydžių skaičius n didesnis ar y_0 ir y_n skirtumas ryškesnis. Jei interpoliuotinų dydžių skaičius n mažas ir, be to, kraštutinių žinomų dydžių skirtumas nedidelis, abudu metodai duos labai artimus skaičius.

3) Interpoliacija vidurkiais.

Jeigu tarp turimųjų y_0 ir y_n dydžių yra reikalinga interpoliuoti tik vienas vidurinis dydis (medianas), praktiškiau ir paprasčiau iš karto interpoliuoti aritmetiniu arba geometriniu vidurkiu sulig tuo, kuris interpoliavimo metodas duotose sąlygose yra statistiškai prasmingesnis. Pav., patiektame pavyzdiniame skaičiavime, kai $y_0 = 100$ ir $y_n = 120$, o $n = 10$, aritmetinis metodas yra reikalingas konstantos *b*, kurią gavus, pats vidurinis interpoliuotasis dydis y_5 pagal tiesės formulą gaunamas:

$$y_5 = a + bx = 100 + 2.5 = 110.$$

Geometrinis metodas, kaip žinoma, pirmiau reikalingas gana komplikuoto skaičiavimo gauti skaičius reiškiniui $(1 + p)$, o jį gavus — dar privalo sprendimo pagal formulą:

$$y_5 = y_0 (1 + p)^5 = 100 (1{,}0185)^5 (=) 109{,}5.$$

Tuo tarpu interpoliacija vidurkiais artutinius skaičius duoda iš karto:

a) aritmetiniu metodu pagal formulą:

$$\bar{y}_a = \frac{y_n + y_0}{2} = \frac{120 + 100}{2} = 110 = y_5;$$

b) geometriniu metodu pagal geometrinių vidurkių formulą XV, kur:

$$\bar{y}_g = \sqrt[n]{y_1 . y_2 . \ldots . y_n} = \sqrt{y_0 y_n} = \sqrt{100 . 120} (=) 109{,}5.$$

Šis interpoliavimo vidurkiais metodas yra tiek paprastas, kad tie, kuriems naudojimas lygčių su pastoviais ir kintamais dydžiais technikinių sunkumų daro, gali pasitenkinti vidurkinės interpoliacijos metodu ir aplamai imant. Šiuo atveju tarpiniai dydžiai tarp y_0 ir jau interpoliuotojo $\bar{y}_a$ arba $\bar{y}_g$ (taip pat tarp $\bar{y}_0$ ir y_n) interpoliuojami vidurkiais tuo pačiu būdu, kaip buvo interpoliuota $\bar{y}_a$ (arba $\bar{y}_g$) tarp y_0 ir y_n. Savaime suprantama, kad interpoliuoti galima ne tik čia apibrėžtais aritmetiniais ir geometriniais, bet ir kvadratiniais, harmoniniais, antiharmoniniais ir kt. vidurkiais.

Pastaroji aplinkybė vad. vidurkinį metodą interpoliacijoms daro ligi tam tikro laipsnio pritaikomu tais atvejais, kai interpoliuotinų dydžių kitimai vyksta neprisilaikydami griežtai nei aritmetinės, nei geometrinės proporcijos, o, tarsime, kurios nors eilės aukštesnio laipsnio parabolinės kreivės kitimų dėsningumu. Pav., jei mūsų panaudotame pavyzdyje kuris nors reiškinys laiko tarpe tarp y_0 ir y_n, kuris, sakysime, tęsėsi 10 metų, pirmąjį periodą maždaug ligi y_3 plėtėsi vis smarkiau ir smarkiau (sakysim, iš karto geometrinėj progresijoj), paskui pamažu perėjo į aritmetinę progresiją, o antrame viso periodo laikotarpyje augimas vyksta tik geometrinės proporcijos (vienodais) koeficientais, tai vidurkiniu metodu šios rūšies kitimus galėsime apytikriai atvaizduoti šitaip. Pirmiausia gauname y_5, kurį atstovaus aritmetinis vidurkis iš y_0 ir y_{10} pagal formulą

$$\bar{y} = \frac{y_0 + y_{10}}{2} = y_5.$$

Po to pirmojo periodo reikalingąjį dydžių kiekį tarp y_0 ir $\bar{y}$ interpoliuojame kvadratiniais vidurkiais pagal formulą XV, o antrojo periodo dydžiai (tarp $\bar{y}$ ir y_{10}) interpoliuojami geometriniais vidurkiais. Tuo būdu gautaisiais skaičiais apibrėžtoji kitimų linkmė maždaug atitiks trečiosios eilės parabolę, kuri interpoliacijoms taip pat gali būti naudojama.

4) Kiti interpoliacijos metodai.

Tiek aritmetinė, tiek geometrinė interpoliacija yra sąlygojama vienodų (aritmetinėj ar geometrinėj proporcijoj) linkmės kitimų. Keičiant vidurkių sistemą vad. vidurkinio interpoliavimo metode galima ir bet kurios linkmės specifines kryptis atvaizduoti. Tačiau šitame atvaizdavime kiekvienu atveju daug nulemia nuožiūra. Berods, patsai interpoliavimas savaime jau yra nuožiūros dalykas, ir vieno ar kito interpoliavimo metodo pasirinkimas gali nulemti bet kurio interpoliuotojo dydžio ar kurio interpoliuotųjų dydžių dalinio lyginamąjį svorį. Aplamai, statistikos praktikoj tiesioginėj prasmėj interpoliacijos ilgesniems laiko periodams, jei kitimų linkmė aplamai nežinoma, kiek galint vengiama. Tačiau tais atvejais, kai kitimų linkmė bent apytikriai yra žinoma, artutiniai dydžiai neperil-

giems periodams duodasi vienu ar kitu metodu interpoliuojami pagal atitinkamas formulas. Žinomiausios iš tų formulų yra Newtono arba Newton'o - Gauss'o interpoliavimo formulos (taip pat Stirling'o, Everett'o ir kt.) variantai, pritaikinti įvairios linkmės bei dispersijos tipams. Susipažinti su šiais interpoliacijos metodais galima iš Rietz - Bauer, Handbuch der mathematischen Statistik, III (Glover, Interpolation..., 47—71 pp.), Westergaard - Nybolle, Grundzüge der st., pp. 318—353, ir kt.

Interpoliuotųjų dydžių tikimybės laipsnis bei vidutinė ir galimoji paklaida, taip pat be kurio dydžio maksimalinė tikimybė gali būti matuojama pagal atitinkamas formulas iš IV dalies 7—9 skyrių (statistinės ir matematinės tikimybės bei galimos paklaidos).

5) Linkmės dalinių interpoliacija.

Pasitaiko, kad turimieji statistikos duomenys apima stambius skaitomuosius periodus, bet nėra duomenų smulkesniems laikotarpiams. Pav., gali būti kiekybinės dydžio klasės ilgais intervalais (10 metinės amžiaus grupės, metiniai užsienių prekybos, biudžeto ar natūrinio gyventojų judėjimo duomenys). O reikalinga gali būti metinės amžiaus grupės, mėnesiniai ar kt. eksporto ar biudžeto pajamų duomenys. Tarp kita ko, pav., Anglijos statistikos šaltiniuose dažnai galima rasti tik kvartalinių (trijų mėnesių) st. duomenų, kai kitų valstybių st. šaltiniuose tos pačios rūšies duomenys skelbiami kas mėnuo. Tokiu atveju, norint gauti reikalingus mėnesinius duomenis lyginamajai statistikai, tenka trūkstamus linkmės dalinius (mėnesinius duomenis) interpoliuoti pasikliaujant turimais ilgesnio laikotarpio duomenimis.

Tarkime, kad mes turime metinius kurios nors prekės eksporto duomenis, o reikalingi yra mėnesiniai duomenys. Žinoma, paprasčiausias ir lengviausias būdas gauti mėnesinius duomenis — paimti metinių duomenų $\frac{1}{12}$ dalį. Tačiau retai pasitaiko, kad kuris nors įvykis kas mėnuo vienodu ir tuo pačiu skaičiumi pasireikštų. Ypač ekonominės (užsienių prekybos,

biudžeto, ir kt.) statistikos visa eilė svarbių reiškinių vyksta bei kinta tam tikrais periodais arba sezonais (pav., javų eksportas didesnis rudenį ir žiemą, pieno produktų eksportas padidėja pavasarį, tam tikri mokesčiai mokami ketvirčių gale ir t. t.). Visais tokiais ir panašiais atsitikimais linkmės dalinių interpoliacija vienodais dydžiais visai prasmės nustotų. Taip pat neturėtų prasmės, sakysime, metinės gyventojų amžiaus grupės, sudarytos tuo būdu, kad iš bendrojo 0—10 metų amžiaus klasės gyventojų skaičiaus kiekvienų metų amžiaus grupei būtų priskirta $\frac{1}{10}$ dalis. Tokiais ar panašiais atvejais tenka jau panaudoti kurį nors Newton'o - Gauss'o ar kt. formulos variantą, kuris kiek galint atitiktų menamo reiškinio linkmės dalinių pasiskaidymą (dispersiją) visumoje (visoje dydžio klasėje, visame laikotarpyj ir t. t.). Šiam reikalui interpoliacijos formų variantai yra pritaikinti normaliai dispersijai, taip pat U ir tiesioginės bei atvirkščiosios J dispersijos tipams. Šių formulų panaudojimas statistikos praktikoj ta forma, kaip jas rašo matematikai, gali atrodyti nepaprastai sudėtingas dalykas.

Pav., Newton'o - Gauss'o formula normalios dispersijos daliniams interpoliuoti, kaip ji paprastai rašoma,

$$y=y_0+x\triangle y_0+\frac{x(x-1)}{1.2}\triangle^2 y_{-1}+\frac{(x+1)x(x-1)}{1.2.3}\triangle^2 y_{-1}+\ldots\ldots,$$

atrodo gana komplikuota. Iš tikrųjų gi, nesunku, palyginus, įsitikinti, kad savo esmėje yra tai binomo formulos tam tikras variantas. O jeigu taip, tai dalinių linkmės normali dispersija duodasi interpoliuojama ir betarpiškai atitinkamos binomo eilutės koeficientais. Pav., jeigu žinome, kad, sakysim, kiaušinių metinis eksportas sudaro 4,8 mil. litų ir daugiausia susikaupia birželio — liepos mėnesiais ir labai mažai eksportuojama gruodžio — sausio mėnesiais, tai metinį eksportą paskaityti atskiriems mėnesiams galima ir pagal binomo koeficientus iš vienuoliktojo laipsnio eilutės, kur

$$(1+1)^{11}=2\,048.$$

Tuo būdu atskiriems mėnesiams gauname:

Lentelė 35.

LINKMĖS DALINIŲ INTERPOLIACIJA PASIKLIAUJANT BINOMO KOEFICIENTŲ DISPERSIJOS NORMA.

Sąlygos:

bendrąjį 4 800 000 skaičių išskaidyti į 12 linkmės dalinių pagal II-jo laipsnių binomo koeficientų eilutės narių proporciją, kur $(a + b)^n = (1 + 1)^{11} = 2\,048$.

---

y_0 (sausis) $= \frac{1 \cdot 4\,800\,000}{2\,048} =$ 2343,75

y_1 (vasaris) $= \frac{11 \cdot 4\,800\,000}{2\,048} =$ 25781,25

y_2 (kovas) $= \frac{55 \cdot 4\,800\,000}{2\,048} =$ 128906,25

y_3 (balandis) $= \frac{165 \cdot 4\,800\,000}{2\,048} =$ 386718,75

y_4 (gegužis) $= \frac{330 \cdot 4\,800\,000}{2\,048} =$ 773437,50

y_5 (birželis) $= \frac{462 \cdot 4\,800\,000}{2\,048} =$ 1082812,50

y_6 (liepos) $= \frac{462 \cdot 4\,800\,000}{2\,048} =$ 1082812,50

y_7 (rugpiūtis) $= \frac{330 \cdot 4\,800\,000}{2\,048} =$ 773437,50

y_8 (rugsėjis) $= \frac{165 \cdot 4\,800\,000}{2\,048} =$ 386718,75

y_9 (spalis) $= \frac{55 \cdot 4\,800\,000}{2\,048} =$ 128906,25

y_{10} (lapkritis) $= \frac{11 \cdot 4\,800\,000}{2\,048} =$ 25781,25

y_{11} (gruodis) $= \frac{1 \cdot 4\,800\,000}{2\,048} =$ 2343,75

Jei būtų pagrindo teigti, kad susikaupimas apie vidurį metų yra silpnesnis ir, tarsime, gruodžio ir sausio mėn. eksportas sudaro apie 4% metinio eksporto arba aplamai didžiausias susikaupimas ne metų viduryje, o pradžioj arba gale, — tektų arba normalios dispersijos koeficientų eilutės narių santykį pakeisti, arba anormalios dispersijos atitinkamo tipo formulą pasirinkti.

6) Linkmės dalinių ekstrapoliacija.

Ekstrapoliacija nėra reikalinga nei specialiai pritaikintų formulų, nei metodų, be tų, kurie naudojami interpoliacijai aplamai imant. Jei mes turime reikalingus duomenis y ir jeigu, tarsime, interpoliuojame tiesės lygtimi trūkstamus ar reikalingus 9 ar 10 dydžius pagal formulą $y = a + bx$, kur $a = y_0$, o $b = \frac{y_n - y_0}{10}$, tai drąsiai galime teigti, kad ekstrapoliuotieji y_{11} ar y_{-1} dydžiai išeis nemažiau patikimi, kaip interpoliuotieji dydžiai y_5 ar y_6. Pav., psl. *293* patiektame pavyzdyje Lietuvos gyventojų skaičiams interpoliuoti, randame:

$$y_0 = a = 2\,175\,000, \quad y_n = 2\,550\,000, \quad b = 26\,786.$$

Interpoliuojant vidurinį dydį y_7 ir ekstrapoliuojant dydžius y_{-1} ir y_{n+1} tiesės lygtimi pagal formulą

$$y = a + bx = 2\,175\,000 + 26\,786x \text{ gauname:}$$
$$y_7 = 2\,175\,000 + 26\,786 \cdot 7 = 2\,362\,502$$
$$y_{-1} = 2\,175\,000 - 26\,786 = 2\,148\,214$$
$$y_{15} = 2\,175\,000 + 26\,786 \cdot 15 = 2\,576\,790.$$

Panašiu būdu dalinių linkmė gali būti ekstrapoliuojama ir pagal bet kurią kitą formulą. Ekstrapoliuotųjų dydžių tikimybės bei paklaidos ribos kaip ir interpoliuotųjų, gali būti apibrėžiamos taip pat XXVIII, XXXIII ir XXXV formulų pagalba.

Aplamai, tačiau, visais tais atvejais, kai yra atitinkamų statistinių duomenų praeitam laikui, užuot tiesioginės interpoliacijos arba ekstrapoliacijos, yra prasmingiau ir tikslingiau reikiamus duomenis gauti pasirėmus koreliacijos koeficientais pagal regresinės formulos vieną ar kitą variantą.

Literatūra ir papildomos žinios: 1) *Rietz-Bauer,* Handbuch der mathemat. St. pp. 47—80 *(J. W. Glover,* Interpolation etc.), 2) *A. Bowley,* Elements of St., pp. 214—241, 3) *L. March,* Les Principes de la Methode Statistique, pp. 475—502, 4) *L. R. Connor,* Statistics, pp. 177—197, ir kt.

11. Koreliacija ir statistinės pareinamybės.

a) Koreliacijos sąvoka ir rūšys.

Koreliacija, kaip patsai žodis rodo, reiškia savitarpio santykį ar santykiavimą. Matematikoj, o taip pat ir statistikoj, koreliacijos sąvoka apima dviejų, ar daugiau kaip dviejų, st. eilučių pareinamybę, kuri vienu ar kitu metodu gali būti apibrėžiama bei atvaizduojama. Pav., stebint gimimų — mirimų, eksporto — importo, biudžeto pajamų — išlaidų ir kt. duomenis, kai kada tam tikri abiejų st. eilučių saitai matomi ir iš akies: auga gimimai — didėja ir mirimai, mažėja eksportas — mažėja ir importas ir t. t. Arba ekonomikoj žinomieji pasiūlų — paklausų, mažėjančio žemės našumo ir kt. vad. ekonomikos dėsniai kaip tik reiškia vienos ar kitos rūšies tam tikrą pareinamybę.

Šios rūšies pareinamybės buvimą ar nebuvimą dažnai išrodinėjama ir statistikos skaičiais, arba sulig tais skaičiais nubrėžtomis kreivėmis — grafiškai. Ir kai savitarpio pareinamybės, kurios gali būti ir grynai formalinės (vad. paralelizmai), pakankamai ryškios — toksai vaizdavimas gali būti pakankamas. Tačiau gali ir taip atsitikti, kad kreivėmis vaizduojami dvejeto ar keletos eilučių statistikos duomenys, kai jie atstovauja nevienareikšmius dydžius (pav., pasiūlai rodomi kiekio matais, o kaina — pinigo vienetais) arba bent kiek nesutampa periodai, susipina ir koreliacijos žymes aptemdo. Tokiais atvejais kai kada mėginama turimus empirinio st. stebėjimo duomenis „išlyginti", indeksų pagalba „suderinti" ir p. Bet ne visuomet ir šios priemonės gelbsti. Be to, visi šios rūšies pareinamybių vaizdavimai geriausiu atveju teparodo koreliacinio ryšio buvimą ar nebuvimą. O gali būti labai svarbu žinoti ir pareinamybės laipsnį, t. y. jos tikimybinį dydį. Šį dydį

gi, t. y. pareinamybės tikimybę (tikrenybės laipsnį), geriausia apibrėžia koreliacijos koeficientas.

Tačiau iš anksto tenka įsidėmėti, kad patsai koreliacijos koeficientas, vistiek kuriuo metodu gautas, savaime vaizduoja tik matematišką arba formalinę dviejų ar daugelio st. eilučių pareinamybę, o ne kauzalinę, t. y. ne materiališkai priežastinę. Priežastinės pareinamybės, kaip ir „priežasčių priežasties" ieškojimas, priklauso jau kitų mokslų, ne statistikos. Statistika, tačiau, ir statistiniai koreliacijos koeficientai žymiai gali palengvinti ir priežastinių pareinamybių esimo nustatymą.

b) Koreliacijos koeficiento pakaitalai.

Tais atvejais, kai norima nustatyti tik dviejų st. eilučių koreliacijos buvimą ar nebuvimą, galima pasikliauti ir Julin'o vadinamojo „pareinamybės indekso" formula[10], kurią jis rašo šitaip:

$$\frac{c-d}{c+d} = \frac{c-d}{n}$$

Simbolis *c* Julin'o formuloj atstovauja pozityvinių reiškinių skaičių, d — negatyvinių ir n — bendrąjį skaičių. Reiškinių charakteristika atžymima + arba — sulig tuo, ar x-inės ir y-inės eilučių sekos dydžiai didėja ar mažėja. Julin'o formulą pritaikant mūsų naudojamai formulų simbolių sistemai, tektų rašyti:

$$P_r = \frac{\Sigma \pm}{n} = \frac{c-d}{n} = \frac{c-d}{|c+d|} \qquad \text{(XL)}.$$

Nes, iš tikrųjų, c ir d Julin'o formuloj atstovauja ne pozityvinius ar negatyvinius skaičius, o tik pozityvinių ar negatyvinių reiškinių (+ ir —) skaičius. Pavyzdžiui, paėmę gimimų ir natūrinio prieauglio bendruosius lyginius, gausime:

---

[10] Žiūr. Julin, Principles de Statistique etc., 499.

Lentelė 36.

GIMIMŲ IR NATŪRINIO PRIEAUGLIO „PAREINAMYBĖS INDEKSAI“ PAGAL JULIN'Ą (formula XL).

Metai	Gimimų lyginiai (‰)		Natūrinio prieauglio lyginiai (‰)		± .
	y	±	x	±	
1922	27,3	(±)	9,6	(±)	(±)
1923	28,4	+	13,3	+	+
1924	29,4	+	13,1	—	—
1925	28,9	—	12,0	—	+
1926	28,5	—	13,1	+	—
1927	29,4	+	12,1	—	—
1928	28,8	—	13,2	+	—
1929	27,2	—	10,1	—	+
1930	27,4	+	11,5	+	+
1931	26,8	—	11,0	—	+
1932	27,3	+	12,0	+	+
1933	25,7	—	12,1	+	—
1934	24,8	—	10,2	—	+
1935	23,4	—	9,4	—	+
1936	24,2	+	10,6	+	+
				$\Sigma\pm = 9 - 5 = 4$	

Tuo būdu Julin'o „pareinamybės indekso“ formula duoda devynis pliusus (+ 9) ir penkis minusus (— 5). Taigi pareinamybės indeksas gaunamas:

$$P_r = \frac{\Sigma\pm}{n} = \frac{c - d}{|c + d|} = \frac{9 - 5}{9 + 5} = \frac{4}{14} (=) 0{,}29.$$

Aplamai, kaip žinome, koreliacijos koeficiento uždavinys — parodyti pareinamybės l a i p s n į. O laipsnis — tai tikimybinis dydis: jis turi parodyti atstumą nuo tikrenybės, kuri atstovauja visumą (t. y. 1 arba 100%) be jokių išimčių. Kai pareinamybės indeksas pareikštas p = 0,29 trupmena, kuri yra artesnė 0 negu 1, pati pareinamybė kiekvienam turi atrodyti mažai patikima. O iš to tektų daryti išvadą, kad tarp gimimų skaičiaus ir natūrinio prieauglio negalima rasti ne tik kauzalinės pareinamybės, bet nė formalinio ryšio arba koreliacijos. O tai jau išvada visai priešinga gyvenimo faktų logikai.

Panašiais atvejais tenka ieškoti „teorijos ir gyvenimo“ nesutapimo priežasčių. O priežastis gali būti arba „teorijoj“ (šiuo atveju formuloj) ar jos netinkamam naudojime. Šiuo konkretišku atveju, sakysim, galima būtų prikišti, kad atsižvelgta pirmųjų skirtumų tendencijos (eilutės sekoj dydžių kitimo), o ne skirtumų nuo vidurkio. Ir, tiesa, tyrinėdami skirtumus nuo vidurkio (žiūr. skilt. 7 lentelė 37), randama + 11 ir — 3. Taigi ta pati formula jau duoda:

$$P_r = \frac{11-3}{11+3} = \frac{9}{14} (=) 0{,}64.$$

Deja, šiuo atveju sunku būtų rasti argumentų už reikalą pirmuosius skirtumus pakeisti skirtumais nuo vidurkio. Svarbiausia, kad, kaip „Socialekonominės statistikos“ p. p. 538 ir 539 yra mano įrodyta, Julin'o formula, pakeitus skaičiavimo metodą, iš vieno kraštutinumo dažnai pereina į kitą, ir nėra priemonių patikrinti, kuris metodas tinkamesnis (rugių derliaus, t. y. pasiūlų, ir kainos pareinamybės indeksas 1923—1931 mt. pirmaisiais skirtumais kalendorinių metų išėjo 0, o sulig skirtumais ūkio metais net 1). Priežastis ta, kad Julin'o formula atsižvelgia vien tik skirtumų pobūdžio ir visai paneigia jų lyginamąjį svorį. O gyvenime, taigi ir jo reiškinius vaizduojančiose statistikos eilutėse, visi kitimai pasireiškia dydžiais, kurių santykis nebūtinai turi būti vienodas. Dėl šios priežasties Julin'o formulos neišgelbsti ir jos variantas, kurį patiekia, tarp kita ko, King'as jo vadinamajai „paprastajai koreliacijai“ atvaizduoti[11].

$$p_r = \sqrt{\frac{2c-n}{n}} \qquad \text{(XLa)}.$$

Patikrinkime šios formulos veikimą tais pačiais pavyzdžiais:

1) Jei c = 9 ir n = 14, kaip lent. 36 King'o „paprastoji koreliacija“ rodo, tai

$$p_r = \sqrt{\frac{2.9-14}{14}} = \sqrt{\frac{18-14}{14}} = \sqrt{\frac{4}{14}} (=) \sqrt{0{,}29} (=) 0{,}53$$

[11] W. I. King, Elements of Statistical Method, 1923, p. 207 ir sek.

2) Jeigu c = 11, n = 14, tai:

$$p_2 = \sqrt{\frac{2 \cdot 11 - 14}{14}} = \sqrt{\frac{22 - 14}{14}} = \sqrt{\frac{8}{14}} (=) \sqrt{0{,}57} (=) 0{,}75.$$

Kaip matome, King'o formula, kaip Julin'o formulos variantas, šią pastarąją sukompromituoja. O jei imti mano nurodytus pavyzdžius „Socialekonominėj Statistikoj" p. 538, kur tie patys duomenys vienu atveju iš n = 8 duoda c = 4, o kitu iš tų pačių n = 8 išeina c = 0 tai gausime:

$$1)\ p_r = \sqrt{\frac{2c - n}{n}} = \sqrt{\frac{2 \cdot 4 - 8}{8}} = \sqrt{\frac{8 - 8}{8}} = 0 \text{ ir}$$

$$2)\ p_r = \sqrt{\frac{2 \cdot 0 - 8}{8}} = \sqrt{\frac{-8}{8}} = \sqrt{-1} = i,$$

t. y. lygu menamajam vienetui.

Kaip matome, čia jau susikompromituoja pati King'o formula. Aplamai, jei statistikoj vienam ar kitam reikalui norima naudoti tikimybiniai dydžiai (o trupmenoj pareiškiama koreliacija privalo rodyti pareinamybės — kauzalinės ar tik formalinės — laipsnį, t. y. santykį su tikrenybe), tai jų suradimas bei skaičiavimo metodai turi atitikti tikimybių teorijos reikalavimus bei tikimybių skaičiuotės dėsnius. Šiuos reikalavimus tam tikromis sąlygomis patenkina vad. P e a r s o n'o k o r e l i a c i j o s k o e f i c i e n t a i.

c) Pearson'o koreliacijos koeficientai.

Pearson'o koreliacijos koeficientas skaičiuojamas pagal formulą:

$$r = \frac{\Sigma \triangle_x \triangle_y}{n\, \sigma_x \sigma_y} = \frac{\frac{1}{n} \Sigma (x - \bar{x})(y - \bar{y})}{\sqrt{\frac{\Sigma \triangle_x^2}{n}} \cdot \sqrt{\frac{\Sigma \triangle_y^2}{n}}} \qquad \text{(XLI).}$$

Savaime suprantama, kad koreliacijos koeficientas gali būti skaičiuojamas ir betarpiškai x_i ir y_i dydžių pagal atitinkamą formulos variantą, kur

$$r = \frac{\Sigma x_i y_i - \bar{x}\bar{y}}{n\,\sigma_x\,\sigma_y} = \frac{\frac{1}{n}\Sigma x_i y_i - \bar{x}\bar{y}}{\sqrt{\frac{1}{n}\Sigma x^2 - \bar{x}^2}\,\sqrt{\frac{1}{n}\Sigma y^2 - \bar{y}^2}} \qquad \text{(XLIa).}$$

Formulos rodo, kad r pareina ne tik nuo x-inės ir y-inės eilučių linkmės savitarpio kitimų santykio, bet ir nuo kiekvienos eilutės narių x_i ir y_i dydžio skirtumų, be to, nuo n, t. y. nuo st. eilučių narių skaičiaus. Tuo būdu Paerson'o formula ne tik atsižvelgia koreliuojamų eilučių narių lyginamojo svorio, bet, jei galima taip pasakyti, turi ir automatišką kontrolės aparatą: nes kai r skaičiuojamas pasikliaujant koreliuojamųjų eilučių viduriniais kvadratingaisiais nukrypimais, tai tuo pačiu iš anksto gaunami rodikliai, sulig kuriais galima spręsti, ar — ir kiek — pasireiškia didžiųjų skaičių dėsnis koreliuojamose eilutėse. Savaime suprantama, kad pirmiau negu ieškoti koreliacijos koeficiento, reikalinga patikrinti abiejų eilučių kvadratingųjų nukrypimų, taip pat vidurkių, patikimumo ribas pasikliaujant anksčiau parodytais statistinio analizio metodais. Jei posirodytų, kad galimos labai didelės paklaidos, ar išeitų labai didelis dispersijos koeficientas, — tai būtų rodiklis, kad nėra prasmės nei koreliacijos koeficiento ieškoti.

Patsai koreliacijos koeficientas gali būti skaičiuojamas ne tik sulig skirtumais nuo vidurkio, bet ir sulig pirmaisiais skirtumais — sulig tuo, kokios rūšies kauzalinėj pareinamybėj koreliuojami reiškiniai vyksta ar vykti gali. Pav., „Socialekonominės statistikos" p. p. 540—544 patiektame koreliacijos koeficiento pavyzdiniame skaičiavime derliaus ir kainų pakitimo pareinamybei, koreliacijos koeficientas sulig skirtumais nuo vidurkio išėjo — 0,60, o sulig pirmaisiais skirtumais net — 0,86. Jei paimsime mūsų jau naudotus gimimų ir natūrinio prieauglio duomenis, tai koreliacijos koeficientui sulig skirtumais nuo vidurkio gausime šitokius reiškinius:

Gimimų (x_i) bendrojo lyginio 1922—1936 mt. vidurkis pagal formulą VIII išeina: $\bar{x}$ (=) 27,2, o natūrinio prieauglio $\bar{y}$ (=) 11,6.

Lentelė 37.

GIMIMŲ IR NATŪRINIO PRIEAUGLIO KORELIACIJOS KOEFICIENTAS SULIG SKIRTUMAIS NUO VIDURKIO.

Nr.	Metai	Gimimai (‰)	Skirtumai nuo vidurkio	Skirtumų kvadratai	Natūrinis prieauglis (‰)	Skirtumai nuo vidurkio	Skirtumų kvadratai	Gimimų ir nat. prieauglio skirtumų sandauga
		x_i	$\triangle_x = x_i - \bar{x}$	$\triangle_x^2$	y_i	$\triangle_y = y_i - \bar{y}$	$\triangle_y^2$	$(\triangle_x \triangle_y)$
	0	1	2	3	4	5	6	7
1	1922	27,3	0,1	0,01	9,6	— 2,0	4,00	— 0,20
2	1923	28,4	1,2	1,44	13,3	1,7	2,89	2,04
3	1924	29,4	2,2	4,84	13,1	1,5	2,25	3,30
4	1925	28,9	1,7	2,89	12,0	0,4	0,16	0,68
5	1926	28,5	1,3	1,69	13,1	1,5	2,25	1,95
6	1927	29,4	2,2	4,84	12,1	0,5	0,25	1,10
7	1928	28,8	1,6	2,56	13,2	1,6	2,56	2,56
8	1929	27,2	0	0	10,1	— 1,5	2,25	0
9	1930	27,4	0,2	0,04	11,5	— 0,1	0,01	— 0,02
10	1931	26,8	— 0,4	0,16	11,0	— 0,6	0,36	0,24
11	1932	27,3	0,1	0,01	12,0	0,4	0,16	0,04
12	1933	25,7	— 1,5	2,25	12,1	0,5	0,25	— 0,75
13	1934	24,8	— 2,4	5,76	10,2	— 1,4	1,96	3,36
14	1935	23,4	— 3,8	14,44	9,4	— 2,2	4,84	8,36
15	1936	24,2	— 3,0	9,00	10,6	— 1,0	1,00	3,00
		$407,5 = \Sigma x_i$		$49,93 = \Sigma \triangle_x^2$	$173,3 = \Sigma y_i$		$25,19 = \Sigma \triangle_y^2$	$25,66 = \Sigma \triangle_x \triangle_y$

Viduriniai kvadratingieji nukrypimai gaunami pagal formulą XXVII:

$$\sigma_x = \sqrt{\frac{49,93}{15}} = \sqrt{3,33}\ (=)\ 1,82,$$

$$\sigma_y = \sqrt{\frac{25,19}{15}}\ \sqrt{1,68}\ (=)\ 1,30.$$

Patikrinę gautųjų st. dydžių patikimumą pagal formulas XXXI ir XXXII, gauname patikimąsias paklaidas (u):

$$u_{\bar{x}} = \pm \frac{0,6745\sigma}{\sqrt{n}} = \frac{0,6745 \cdot 1,82}{\sqrt{15}}\ (=)\ \frac{1,23}{3,87}\ (=)\ 0,32;$$

$$u_{\bar{y}} = \frac{0,6745 \cdot 1,3}{\sqrt{15}}\ (=)\ \frac{0,88}{3,87}\ (=)\ 0,23;$$

$$u_{\sigma,x} = \frac{0,6745\sigma}{\sqrt{2 \cdot n}} = \frac{0,6745 \cdot 1,82}{\sqrt{2 \cdot 15}}\ (=)\ \frac{1,23}{\sqrt{30}}\ (=)$$

$$(=)\ \frac{1,23}{5,48}\ (=)\ 0,22;$$

$$u_{\sigma,y} = \frac{0,6745 \cdot 1,3}{\sqrt{30}}\ (=)\ \frac{0,88}{5,48}\ (=)\ 0,16.$$

Tuo būdu randame, kad vidurkių patikimosios paklaidos sudaro $1-2\%$ $\left(\frac{0,32}{27,2}\ (=)\ 0,01 \text{ ir } \frac{0,23}{11,60}\ (=)\ 0,02\right)$, o kvadratingųjų nukrypimų — $2-1\%$ $\left(\frac{0,2}{1,12}\ (=)\ 0,01 \text{ ir } \frac{0,01}{1,32}\ (=)\ 0,01\right)$.

Šiose aplinkybėse visi dydžiai dar galima laikyti patikimais (paprastai patikimoji paklaida gali būti laikoma statistikoj neperdidelė, jei ji neprašoka 3%, o galimoji paklaida — 10%).

Atsargumo dėliai patikrinkime dar ir abiejų eilučių dispersiją pagal formulą XXIVa. Tuo būdu gauname:

$$d_{\bar{x}} = \frac{\sigma}{\bar{x}} = \frac{1,82}{27,2}\ (=)\ 0,066 \text{ arba apie } 7\% \text{ ir}$$

$$d_y = \frac{1,3}{11,6} (=) 0,11 \text{ arba apie } 11\%.$$

Tuo tarpu žinome, kad tik tuo atveju, kai dispersijos koeficientas pasiekia 33% vidurkio ir artinasi prie 50%, yra pagrindo abejoti dispersijos normalumu, jei tuo pačiu laiku n yra pakankamai didelis skaičius.

Taigi turimųjų duomenų analizis rodo, kad abiejose st. eilutėse didžiųjų skaičių dėsnio veikimas yra pasireiškęs ir dėl to koreliacijos koeficiento skaičiavimas gali būti statistiškai prasmingas.

Kai viduriniai kvadratingieji nukrypimai buvo skaičiuoti pagal skirtumus nuo vidurkio, tai, turint skirtumus $\triangle_x$ ir $\triangle_y$, patogiausias pirmasis XLI formulos variantas, kur

$$r = \frac{\Sigma \triangle_x \triangle_y}{n \sigma_x \sigma_y}.$$

Iš lentelės 37, skilties 7 duomenų matome, kad abiejų eilučių skirtumų sandaugos algebrinė suma $\Sigma \triangle_x \triangle_y = 25,66$. Taigi pagal formulą XLI

$$r = \frac{25,66}{15 . 1,82 . 1,3} = \frac{25,66}{35,49} (=) 0,72.$$

Šitaip pareikštas koreliacijos koeficientas duoda pagrindo teigti, kad natūrinio prieauglio tiesioginė pareinamybė nuo gimimų turi apie 72% tikimybės, o apie 28% tenka kitiems veiksniams.

Kai r = 0,72, taigi skirtumas nuo tikrenybės tik 0,28, yra pagrindo teigti, kad r turinio esmės negali pakeisti ir patikimoji paklaida. Vis dėlto tais atsitikimais, kai n nėra labai didelis skaičius, visuomet pravartu r apibrėžti dar ir patikimąja paklaida, panašiai kaip apibrėžėme $\bar{x}$, $\bar{y}$ ir σ patikimąsias paklaidas. Tam reikalui galioja ši formula:

$$u_r = \frac{1 - r^2}{\sqrt{n}} 0,6745 \qquad \text{(XLII)}.$$

Tuo būdu gauname:

$$u_r = \frac{1 - 0,72^2}{\sqrt{15}} 0,6745 (=) \frac{1 - 0,52}{3,87} = \frac{0,48}{3,87} (=) 0,12.$$

Taigi randame, kad net blogiausiu atveju koreliacijos koeficientas prašoka 60%. Nes kai $r \pm u$, o $r = 0{,}72$ ir $u = 0{,}12$, tai $0{,}72 \pm 0{,}12$ sudaro intervalą $r = 0{,}60 \leftrightarrow 0{,}84$. Jeigu n būtų ne 15, o dvigubai didesnis skaičius (30), tai u tesudarytų tik $\frac{0{,}48}{5{,}48} (=) 0{,}09$. O jeigu $n = 100$, tai $u = \frac{0{,}48}{10} (=) 0{,}048$, ir tuo būdu praktiškai koreliacijos koeficiento jau nebeveiktų. Šiuo pavyzdžiu dar kartą pabrėžiamas n skaičiaus svarbumas visuose statistikos, o ypač statistinių tikimybių, skaičiavimuose.

Koreliacijos koeficientas sulig pirmaisiais skirtumais turi statistinę prasmę tais atvejais, kai yra pagrindo tarti, kad dviejų st. eilučių pareinamybę nulemiantys veiksniai pasireiškia ne betarpiškai, o kuriam laikui praėjus. Pati skaičiavimo technika savo esmėje ta pati, tik skirtumai imami ne nuo vidurkio, o nuo pirmesniųjų eilutės dydžių. Kai mūsų simbolių sistemoj esame sutarę žymėti skirtumus nuo vidurkio $x_i - \bar{x} = \triangle_x$, tai sutarkime žymėti ir pirmuosius skirtumus $x_i - x_{i-1} = \delta_x$. Pavyzdiniam skaičiavimui paimkime tuos pačius duomenis (pusl. 314).

Iš lentelės 38 duomenų koreliacijos koeficientui gauname šiuos reiškinius:

1. $\sigma_x = \sqrt{\frac{\Sigma \delta_x^2}{n}} (=) \sqrt{\frac{12{,}95}{15}} (=) \sqrt{0{,}86} (=) 0{,}93.$

2. $\sigma_y = \sqrt{\frac{\Sigma \delta_y^2}{n}} (=) \sqrt{\frac{36{,}88}{15}} (=) \sqrt{2{,}46} (=) 1{,}57.$

3. $\Sigma \delta_x \delta_y \quad (=) 12{,}09.$

Gimimų ir natūrinio prieauglio koreliacijos koeficientas sulig pirmaisiais skirtumais gaunamas pagal formulą XLIb:

$$r = \frac{\Sigma \delta_{\bar{x}} \delta_y}{n \sigma_{\bar{x}} \sigma_y} = \frac{12{,}09}{15 \cdot 0{,}93 \cdot 1{,}57} (=) \frac{12{,}09}{21{,}85} (=) 0{,}55.$$

Šiuo atveju, kaip matome, koreliacijos koeficientas sulig pirmaisiais skirtumais išeina žymiai mažesnis už koreliacijos koeficientą sulig skirtumais nuo vidurkio. Savaime, tačiau, tai dar nėra įrodymas, kad kitu metodu gautas mažesnis korelia-

Lentelė 38.

GIMIMŲ IR NATŪRINIŲ PRIEAUGLIO KORELIACIJOS KOEFICIENTAS SULIG PIRMAISIAIS SKIRTUMAIS.

Nr.	Metai	Gimimai (‰)	Pirmieji skirtumai	Skirtumų kvadratai	Natūrinis prieauglis (‰)	Pirmieji skirtumai	Skirtumų kvadratai	Gim. ir natūr. prieauglio skirtumų sandauga
		x_i	$\delta_x = x_i - x_{i-1}$	δ_x^2	y_i	$\delta_y = y_i = y_{i-1}$	δ_y^2	$(\delta_x \cdot \delta_y)$
	0	1	2	3	4	5	6	7
1	1922	27,3	0	0	9,6	0	0	0
2	1923	28,4	1,1	1,21	13,3	3,7	13,69	4,07
3	1924	29,4	1,0	1,00	13,1	— 0,2	0,04	— 0,20
4	1925	28,9	— 0,5	0,25	12,0	— 1,1	1,21	0,55
5	1926	28,5	— 0,4	0,16	13,1	1,1	1,21	— 0,44
6	1927	29,4	0,9	0,81	12,1	— 1,0	1,00	— 0,90
7	1928	28,8	— 0,6	0,36	13,2	1,1	1,21	— 0,66
8	1929	27,2	— 1,6	2,56	10,1	— 3,1	9,61	4,96
9	1930	27,4	0,2	0,04	11,5	1,4	1,96	0,28
10	1931	26,8	— 0,6	0,36	11,0	— 0,5	0,25	0,30
11	1932	27,3	0,5	0,25	12,0	1,0	1,00	0,50
12	1933	25,7	— 1,6	2,56	12,1	0,1	0,01	— 0,16
13	1934	24,8	— 0,9	0,81	10,2	— 1,9	3,61	1,71
14	1935	23,4	— 1,4	1,96	9,4	— 0,8	0,64	1,12
15	1936	24,2	0,8	0,64	10,6	1,2	1,44	0,96
		$407,5 = \Sigma x_i$		$12,97 = \Sigma \delta_x^2$	$173,3 = \Sigma y_i$		$36,88 = \Sigma \delta_y^2$	$12,09 = \Sigma \delta_x \delta_y$

cijos koeficientas yra mažiau prasmingas. Bet kuriuo metodu gautojo kor. koeficiento st. prasmingumas pareina nuo koreliuojamų st. eilučių kitimo dėsningumų. Ir vienas ir kitas metodas naudotinas ne savivalingai ar atsitiktinai, o po visapusiško ir kruopštaus koreliuojamų st. eilučių kitimo dėsningumų išnagrinėjimo.

Pearson'o koreliacijos koeficientas tuo tarpu laikomas patikimiausiu pareinamybės rodikliu, nors socialekonominio gyvenimo gana komplikuotus reiškinius atvaizduoti ir jis nevisuomet įstengia. Svarbiausias statistikos praktikų daromas priekaištas[13] tai tas, kad Pearson'o metodu skaičiuojant gaunamas labai mažas koreliacijos koeficientas tuo atveju, kai pagrindinė koreliuojamų eilučių linkmė sutampa, tačiau gretimųjų dydžių skirtumai turi priešingus pažymius. Tačiau šis priekaištas nėra esminis, nes tokiais ar panašiais atvejais gali ne tiek būti kaltas Pearson'o metodas, kiek skaičiuotojo nepakankamas statistikos duomenų paruošimas koreliacijos koeficientui skaičiuoti. Jeigu pagrindinė kitimų linkmė sutampa, o nesutampa tik koreliuojamųjų st. eilučių gretimųjų kitimų reiškiniai, tai yra viena iš dviejų: a) arba koreliuojami reiškiniai yra betarpiškai atvirkščioj pareinamybėj, kuri pasireiškia tik tam tikram laikui praėjus (pav., pasiūlų ir kainos kitimai) ir tokiu atveju turi būti gautas aukšto laipsnio koreliacijos koeficientas su neigiamu pažymiu, arba — b), jei dalykų logika rodo, kad koreliacinė pareinamybė tegali būti tik tiesioginė (pav., gimimų, mirimų ir natūrinio prieauglio), — tad gretimųjų st. eilučių kitimų nesutapimai rodys, kad vienas kuris reiškinys laiko bėgyje kiek atsilieka; tokiu atveju yra reikalinga, pirmiau negu koreliacijos koeficientą vienu ar kitu metodu skaičiuoti, šituos įvykių eigą neatitinkančių statistinių periodų sudarytus skirtumus eliminuoti (st. eilutes išlyginti, skaitomuosius periodus suderinti atsižvelgiant kitimų eigos ir tempo, koreliacijos koeficientą skaičiuoti iš skirtumų nuo pagrindinės linkmės etc.).

d) Kolektyvinis koreliacijos koeficientas.

Kolektyviniu koreliacijos koeficientu mes sutarsime vadinti bendrąjį, vienu ar kitu metodu išskaičiuotą, koreliacijos

[13] Pal. *F. H. Harper,* Elements of Statistics, N. Y. 1930, psl. 208 ir kt.

koeficientą daugiau kaip dviejų st. eilučių savitarpio pareinamybei (koreliacijai) aprėžti. Pav., iš anksčiau patiektųjų pavyzdžių matyti, kad natūrinio prieauglio ir gimimų betarpiškas koreliacijos koeficientas tesudaro apie 0,72 arba apie 72% tikimybės, taigi dar apie 0,28 arba 28% yra nutolęs nuo tikrenybės. Vadinasi, be gimimų, yra dar ir kitų veiksnių, kurie natūriniam prieaugliui daro įtaką. Jeigu natūrinis prieauglis pareitų vien tik nuo gimimų skaičiaus, tai būtų $p = r = 1$. Tačiau mes nustatėme, kad natūrinio prieauglio ir gimimų koreliacijos koeficientas $r = 0{,}72$. O tai reiškia, kad natūrinis prieauglis, matyt, pareina ir nuo kitų priežasčių. Pagal tikimybių sudėties dėsnį, pozityviškai žinoma tikimybė p su nežinomąja arba negalimąja q sudaro sumą, lygią vienetui $p + q = 1$, ir $q = 1 - p$. O kai mūsų tyrinėjime $p = r = 0{,}72$, tai iš to seka, kad

$$q = 1 - r = 1 - 0{,}72 = 0{,}28.$$

Klausimas kas nulemia ligi tikrenybės (1) trūkstamą trupmeną q? Kadangi natūrinis prieauglis, kaip žinoma, skaičiuojamas pagal formulą

natūr. prieauglis = gimimai — mirimai

tai prašyte prašosi mintis, jog natūrinio prieauglio koreliacijos koeficientui ligi vieneto trūkstamą skirtumą nulemia mirimai. Ir ši mintis savo esmėje yra teisinga.

Lentelėj 39 yra patiekti pirmiausia mirimų duomenys, panašiai sutvarkyti kaip gimimų ir natūrinio prieauglio duomenys lentelėj 37. Penkiolikos metų mirtingumo lyginių vidurkis sudaro apie 15,6‰, o vidurinis kvadratingasis nukrypimas:

$$\sigma_2 = \sqrt{\frac{\Sigma \triangle_z^2}{n}} = \sqrt{\frac{25{,}05}{15}} = \sqrt{1{,}67} \quad (=) \ 1{,}29.$$

Toliau skiltyj 4 yra patiekti natūrinio prieauglio skirtumai nuo vidurkio, sk. 5 — gimimų skirtumai nuo vidurkio, sk. 6 — gimimų ir natūrinio prieauglio skirtumų sandaugos ($\triangle_x \triangle_y$), sk. 7 — gimimų ir mirimų skirtumų sandaugos ($\triangle_x \triangle_z$) ir sk. 8 — mirimų ir natūrinio prieauglio skirtumų sandaugos ($\triangle_z \triangle_y$). Iš anksčiau (psl. 311) žinome, kad gimimų kvadratingasis nukrypimas $\sigma_{\bar{x}} = 1{,}82$, natūrinio prieauglio —

Lentelė 39.

SCHEMATINIAI DUOMENYS GIMIMŲ, MIRIMŲ IR NATŪRINIO PRIEAUGLIO KORELIACIJOS KOEFICIENTAMS

Nr.	Metai	Mirimų st. duomenys			Nat. prieauglio skirtumai nuo vidurkio (žiūr. lent. 37, skilt. 5)	Gimimų skirtumai nuo vidurkio (lent. 37, sk. 2)	Skirtumų sandaugos		
		Mirimų lyginiai (‰)	Skirtumai nuo vidurkio	Skirtumų kvadratai			Gimimų su nat. prieauglio skirtumų sandauga	Gimimų su mirimų skirtumų sandauga	Natūrinio prieauglio su mirimų skirt. sandauga
		z_i	$z_i - \bar{z} = \triangle_z$	$\triangle_z^2$	$\triangle_y = y_i - \bar{y}$	$\triangle_{\acute{x}} = x_i - \bar{x}$	$\triangle_{\acute{x}} \triangle_y$	$\triangle_{\acute{x}} \triangle_z$	$\triangle_y \triangle_z$
n	0	1	2	3	4	5	6	7	8
1	1922	17,7	2,1	4,41	— 2,0	0,1	— 0,20	0,21	— 4,20
2	1923	15,1	— 0,5	0,25	1,7	1,2	2,04	— 0,60	— 0,85
3	1924	16,3	0,7	0,49	1,5	2,2	3,30	1,54	1,05
4	1925	16,9	1,3	1,69	0,4	1,7	0,68	2,21	0,52
5	1926	15,4	— 0,2	0,04	1,5	1,3	1,95	— 0,26	— 0,30
6	1927	17,3	1,7	2,89	0,5	2,2	1,10	3,74	0,85
7	1928	15,6	0	0	1,6	1,6	2,56	0	0
8	1929	17,1	1,5	2,25	— 1,5	0	0	0	— 2,25
9	1930	15,9	0,3	0,09	— 0,1	0,2	— 0,02	0,06	— 0,03
10	1931	15,8	0,2	0,04	— 0,6	— 0,4	0,24	— 0,08	— 0,12
11	1932	15,3	— 0,3	0,09	0,4	0,1	0,04	— 0,03	— 0,12
12	1933	13,5	— 2,1	4,41	0,5	— 1,5	— 0,75	3,15	— 1,05
13	1934	14,6	— 1,0	1,00	— 1,4	— 2,4	3,36	2,40	1,40
14	1935	14,0	— 1,6	2,56	— 2,2	— 3,8	8,36	6,08	3,52
15	1936	13,4	— 2,2	4,84	— 1,0	— 3,0	3,00	6,60	2,20
		$233,9 = \Sigma z_i$		$25,05 = \Sigma \triangle_z^2$			$25,66 = \Sigma \triangle_x \triangle_y$	$25,02 = \Sigma \triangle_x \triangle_3$	$0,62 = \triangle \Sigma_y \triangle_3$

$\sigma_y = 1{,}30$. Taigi mums rūpimas mirimų ir natūrinio prieauglio koreliacijos koeficientas pagal formulą XLI gaunamas:

$$r_{zy} = \frac{\Sigma \triangle_z \triangle_y}{n\,\sigma_z\,\sigma_y} = \frac{0{,}62}{15 \,.\, 1{,}29 \,.\, 1{,}30} (=) \frac{0{,}62}{25{,}2} (=) 0{,}025.$$

Kaip matome, mirimų ir natūrinio prieauglio koreliacijos koeficientas tiek mažas, kad, lyginant su natūriniu prieaugliu ir gimimų koeficientu ($r_{xy} = 0{,}72$), atrodo net dėmesin neimtinas. Iš kitos, tačiau, pusės gimimų ir mirimų koreliacijos koeficientas išeina beveik toks pat, kaip gimimų ir natūrinio prieauglio:

$$r_{xz} = \frac{\Sigma \triangle_x \triangle_z}{n\,\sigma_x\,\sigma_z} = \frac{25{,}02}{15 \,.\, 1{,}82 \,.\, 1{,}29} = \frac{25{,}02}{35{,}217} (=) 0{,}71.$$

Taigi mirimų ir natūrinio gyventojų judėjimo saitai šiuo gana aukštu koreliacijos koeficientu konstatuojami. Tik tie saitai yra tampresni tarp gimimų ir mirimų, negu tarp mirimų ir natūrinio prieauglio (nulemia kūdikių mirtingumas?) Jeigu taip, tai jau pravartu paieškoti ir betarpiško ryšio tarp visų trijų reiškinių per vad. kolektyvinį koreliacijos koeficientą.

Kaip betarpiškai individualinis, taip ir kolektyvinis koreliacijos koeficientas skaičiuojamas tikimybių daugybos dėsniu pasikliaujant. Kai paprastąjį (individualinį) koreliacijos koeficientą esamę sutarę žymėti mažąja r (taigi $r_{xy} = r_1$, $r_{xz} = r_2$, $r_{zy} = r_3$), kolektyvinį koreliacijos koeficientą galime žymėti R. Tuo būdu visų trijų (x_i, y_i, z_i) st. eilučių kolektyviniam indeksui formulą galime rašyti šitaip:

$$R = \frac{r_{xy} - r_{xz} \,.\, r_{yz}}{\sqrt{(1 - r^2_{xz})\ (1 - r^2_{yz})}} = \frac{r_1 - r_2 \,.\, r_3}{\sqrt{(1 - r_2^2)\ (1 - r_3^2)}} \qquad \text{(XLIII)}.$$

Formula rodo, kad pirmiau reikia gauti individualinius koreliacijos koeficientus r_i betarpiškai visas tris eilutes suporavus. Turimieji duomenys, kaip matome, davė šitokius r_i:
$r_{xy} = 0{,}72$, $r_{xz} = 0{,}025$ ir $r_{yz} = 0{,}71$.

Tuo būdu gauname:

$$R = \frac{r_{xy} - r_{xz} \,.\, r_{yz}}{\sqrt{(1 - r^2_{xz})\ (1 - r^2_{yz})}} = \frac{0{,}72 - 0{,}025 \,.\, 0{,}71}{\sqrt{(1 - 0{,}025^2) \,.\, (1 - 0{,}71^2)}} =$$

$$(=)\frac{0{,}72-0{,}02}{\sqrt{(1-0{,}5041)\,.\,(1-0{,}0063)}}(=)\frac{0{,}70}{\sqrt{0{,}4959\,.\,0{,}9937}}=$$

$$(=)\frac{0{,}70}{\sqrt{0{,}4928}}(=)\frac{0{,}70}{0{,}702}(=)\,0{,}997.$$

Tuo būdu randame, kad gimimų, mirimų ir natūrinio prieauglio kolektyvinis koreliacijos koeficientas R = 0,997, t. y. visai prisiartina vienetui, arba tikrenybei. Kadangi natūrinis prieauglis iš tikrųjų yra absoliučioj funkcinėj pareinamybėj nuo gimimų ir mirimų santykio, tai gautasis 1 — R skirtumas gali pareiti tik nuo skaičiavimo paklaidos (dėl neišvengiamo artutinio skaičiavimo arba nuo to, kad r_i ir R skaičiuota iš natūrinio gyventojų judėjimo lyginių). Kitokiomis sąlygomis (kur koreliuojami reiškiniai komplikuotesni arba pareina nuo daugiau kaip dviejų priežasčių), gali prireikti kolektyvinį koreliacijos koeficientą iš didesnio skaičiaus individualinių koreliacijos koeficientų skaičiuoti. Bet savo esmėje visa tai bus tik jau mums pažįstami skaičiavimo metodų variantai. Tie variantai randami bet kuriame naujesniame matematikos statistikos vadovėlyje. Šituose, tačiau, vadovėliuose paprastai nepaaiškinama, kuris iš turimų tiesinės koreliacijos koeficientų r privalo būti žymimas r_1. O nuo jo daugiausia pareina ir R. Praktikoj gi r_1 imamas didžiausiasis r.

Literatūra ir papildomos žinios apie koreliacijas randama net ir nematematinės statistikos vadovėliuose. Iš vertingesniųjų paminėtina: 1) *O. N. Anderson,* Einführung in die mathematische Statistik, p. 267—307), 2) *Rietz-Bauer,* Handbuch der mathematischen Statistik, p. p. 159—217 (H. L. Rietz ir A. R. Crathorme, taip pat T. L. Kelley, ir W. M. Persons'o rašiniai), 3) *A. L. Bowley,* Elements of Statistics, p. p. 350—433, 4) *G. U. Yule,* An Introduction to the Theorie of Statistics, p. p. 157—253 ir 317—334, 5) *L. March,* Les Principles de la Methode Statistique, p. p. 575—616, 6) *V. J. Romanovskyj,* Elementarnyj kurs matemat. statistiki, p. p. 224—297.

12. Linkmės kitimų tikimybė ir hipotetiniai dydžiai.

a) Linkmės kitimų statistinė prigimtis.

Kai dviejų, ar daugiau kaip dviejų, st. eilučių savitarpio santykį apibrėžiame vienos ar kitos rūšies koreliacijos koeficientu, tariame, kad koreliacijos koeficientas rodo st. eilučių pa-

reinamybės laipsnį. Jei ši pareinamybė yra tik formalinė, tariame, kad y ir x atstovaujamų reiškinių savitarpio pareinamybę apibrėžia tikimybė r. Šiuo atveju koreliacijos koeficientas terodo tik vad. statistinį paralelizmą, kurio priežastis (ar priežastys), t. y. logiškai kauzalinė pareinamybė gali būti susijusi su visai kitos rūšies reiškiniais. Jei gi koreliacijos koeficiento r atstovaujama tikimybė yra ne tik formali, bet ir logiškai kauzalinė, tariame, kad y-inės st. eilutės atstovaujamo reiškinio kitimai nuo x-inės st. eilutės atstovaujamo reiškinio kitimų (arba x nuo y) pareina tiek, kiek sudaro koreliacijos koeficientas r lyginant su 1; skirtumas iš 1 — r parodo, kiek y (ar x) st. eilutės atstovaujamo reiškinio kitimai pareina nuo kitų (z, v etc.) veiksnių įtakos.

Pavyzdžiui galima imti iš vienos pusės importo ir eksporto arba biudžeto pajamų ir išlaidų duomenis, o iš kitos — jau mūsų nauduotuosius gimimų ir natūrinio prieauglio arba, kaip „Socialekonominės Statistikos" psl. 538—543 parodyta, rugių derliaus ir kainos kitimų duomenis. Visais keturiais atvejais gaunami gana dideli koreliacijos koeficientai, tačiau jų statistinė prigimtis nevienoda. Importo ir eksporto arba biudžeto pajamų ir išlaidų koreliacijos koeficientų statistinė prigimtis yra formalinė, nes tiek biudžeto, tiek užsienių prekybos apyvartos kitimai yra nulemiami pirmoj eilėj kiekvieno krašto ūkio struktūros, konjunktūros ir, be to, ekonominės politikos vartojamų priemonių. Struktūriniai bei konjunktūriniai veiksniai ir ekonominės politikos priemonės paprastai veikia visą tautos ūkio reiškinių visumą, — kad ir ne tuo pačiu laiku ar ne tuo pačiu tempu atskiruose ūkio sektoriuose. Dėl to ir užsienių prekybos arba biudžeto statistikos duomenų kitimai tėra tik statistiniai paralelizmai, kurių logiškai kauzalinė pareinamybė, t. y. pagrindinės priežastys, glūdi ūkio struktūros, konjunktūros ar ekonominės politikos priemonių kitimuose. Kitokia, tačiau, statistinė prigimtis pasireiškia gimimų ir natūrinio prieauglio, arba derlių ir kainos kitimų linkmėje. Tegu mes nežinome, dėlko gimimai arba mirimai mažėja ar daugėja, bet tikrai žinome, kad duotojo meto gimimų nei mirimų skaičiai negali pareiti nuo to paties meto natūrinio prieauglio, kurį visuomet nulemia tik gimimų ir mirimų santykis, o ne atvirkščiai. Arba tegu bet kurio javo pakilusios ar nukritusios kainos veikia ir

derlius tuo būdu, kad ūkininkus paskatina sekančiais metais didinti arba mažinti pasėlių plotus. Tačiau duotojo meto kainos tegali būti padarinys tik iš praeitojo derliaus pasiūlų ir einamojo meto paklausų santykio. O tai reiškia, kad bet kurio javo kainos kauzalinė pareinamybė logiškai tegali būti siejama pirmoj eilėj su praeitojo arba einamojo meto derliais.

Ši iš esmės skirtinga linkmės kitimų prigimtis gali būti vaizduojama tais pačiais statistinio analizio metodais. Nes ar tam tikros rūšies masiniai reiškiniai dėsningai koreliuoja dėl to, kad visi jie iššaukiami bei veikiami bendrųjų veiksnių ar vienodos kilmės priežasčių, arba dėl to, kad yra betarpiškoj kauzalinėj pareinamybėj, — statistiškai matematinio analizio metodų atžvilgiu visais atvejais sąlygos tos pačios: tiek žinomoji, tiek tariamoji pareinamybė tegali būti vaizduojama tik koreliacinės pareinamybės dydžiais. Skirtumas tik tas, kad formalinės pareinamybės dydžiai yra ir palieka tik statistinių eilučių paralelizmo rodikliai, kurie nepasako, nei kurios to ar kito paralelinio reiškinio priežastys, nei kuris reiškinys nuo kurio daugiau pareina. Tuo tarpu logiškoji pareinamybė jau parodo kauzalinį ryšį tarp koreliuojamųjų dydžių, atseit, jų betarpiškos pareinamybės laipsnį.

Šitai žinodami galime ir koreliacinėj (formalinėj ar kauzalinėj) pareinamybėj esančių reiškinių linkmės kitimų tikimybes aprėžti.

b) Linkmės kitimų hipotetiniai dydžiai.

Bet kurio reiškinio linkmės kitimų statistiškai hipotetiniai dydžiai gali būti nustatomi pasikliaujant statistikos duomenimis pagal regresijos formulą, kuri paprastai rašoma kaip ir tiesės lygtys:

$$y = a + bx.$$

Kai linkmės kitimų hipotetiniai dydžiai skaičiuojami iš koreliacijos koeficientų, tai a atstovauja y-inės eilutės aritmetinį vidurkį x — x-inės eilutės skirtumas nuo vidurkio, o konstanta b gaunama iš koreliuojamų st. eilučių kvadratingųjų nukrypimų dalmens, padauginto koreliacijos koeficientu pagal formulą:

$$b = r \frac{\sigma_y}{\sigma_x}.$$

Tuo būdu y-inės eilutės hipotetiniams didžiams gauti regresijos formulą galima ir šitaip rašyti:

$$y' = a + bx = \bar{y} + \frac{r\,\sigma_y}{\sigma_x} \cdot \Delta x \qquad \text{(XLIV)}.$$

Jeigu koreliacijos koeficientas buvo gautas sulig pirmaisiais skirtumais, tai ta pati formula gaus šitokį variantą:

$$y' = \bar{y} + r\,\frac{\sigma_y}{\sigma_x}\,\delta y \qquad \text{(XLIV a)}.$$

Jei būtų norima hipotetinius dydžius apibrėžti ne y-inei, o x-inei eilutei, tai reiškinius tektų sukeisti ir formulą XLIV rašytume:

$$x' = a + by = \bar{x} + \frac{r\,\sigma_x}{\sigma_y}\,\Delta_y.$$

Pavyzdžiui paimkime mūsų jau naudotus gimimų ir natūrinio prieauglio duomenis iš lentelės 37 psl. 310, kur x_j atstovauja atskirų metų gimimus, o y_j — natūrinį prieauglį. Kadangi šiuo atveju yra visai aiški logiškai kauzalinė pareinamybė natūrinio prieauglio (y_j) nuo gimimų (x_j), o ne atvirkščiai, tai statistinę prasmę gali turėti tik natūrinio prieauglio hipotetiniai dydžiai. Regresijos formulai reikalingieji skaičiai yra šie (žiūr. psl. 311—312)):

1) natūrinio prieauglio (y-inės eilutės) a. vidurkis $\bar{y}$ (=) 11,6;
2) natūrinio prieauglio kvadratingasis nukrypimas σ_y (=) 1,30;
3) gimimų (x-inės eilutės) kvadr. nukrypimas σ_x (=) 1,82;
4) abiejų eilučių koreliacijos koeficientas r (=) 0,72.

Taigi gauname:

$$a = \bar{y} = 11{,}6 \quad \text{ir}$$

$$b = r\,\frac{\sigma_y}{\sigma_x} = \frac{0{,}72 \,.\, 1{,}30}{1{,}82}\ (=)\ 0{,}5.$$

Tuo būdu hipotetiniams natūrinio prieauglio dydžiams gauti regresijos formula:

$$y' = a + bx = \bar{y} + \frac{r\,\sigma_y}{\sigma_x}\,\triangle_x,$$

pakeitus konstantas skaičiais, duoda:

$$y' = 11{,}6 + 0{,}5\,x = 11{,}6 + 0{,}5\,\triangle_x.$$

Kai kintamuosius x_i dydžius atstovauja gimimų skirtumai nuo vidurkio ($\triangle x = x_i - \bar{x}$), tai galime juos tiesiog išrašyti iš *37* lentelės *2* skilties.

Lentelėj 40 (psl. *324*) natūrinio prieauglio hipotetiniai dydžiai surašyti *2* skiltyj. Gretimoj skiltyj (*3*) patiekti natūrinio prieauglio skaičiai iš statistinio stebėjimo. Skiltyj 4 duodami skirtumai tarp hipotetinių ir statistinių dydžių. Skiltyj 5 — skirtumų kvadratai. Kad atskirti hipotetinius dydžius nuo statistinių, sutarsime juos rašyti y'.

Lentelės 5 skiltyj matyti, kad patiektųjų hipotetinių dydžių skirtumai nuo tikrųjų skaičių, paėmus juos kvadrate ir susumavus, sudaro $\Sigma(y-y')^2 = 12{,}28$. Tuo būdu hipotetinių dydžių kvadratingasis nukrypimas pagal formulą XXVII išeina:

$$\sigma_y = \sqrt{\frac{\Sigma(y-y')^2}{n}} = \sqrt{\frac{12{,}28}{15}}\ (=)\ \sqrt{0{,}82}\ (=)\ 0{,}9.$$

Tuo tarpu originalios y_i skaičių eilutės kvadratingasis nukrypimas $\sigma_y = 1{,}3$, t. y. visa 0,4 arba net 30% didesnis. Dispersijos koeficientas iš hipotetinių dydžių eilutės y tesudaro tik $\frac{0{,}9}{11{,}60}\ (=)\ 0{,}077$ arba apie 7,7%. Taigi hipotetinių dydžių eilutės vidutinė arba „normalioji“ paklaida aprėžiama 7,7%, o galimoji paklaida (trijų sigmų ribose), aprėžiama *23*%. Vadinasi, su 99,7% tikimybe galima teigti, kad jeigu bet kuriais metais gimimų skaičiaus nukrypimas nuo vidurkio (padidės ar sumažės) 10%, tai padidės ar sumažės ir natūrinis prieauglis su nedidesne kaip *23*% paklaida.

Pamėginkime šį teigimą patikrinti. Štai 1937 mt. iš viso buvo 56 393, ir tai sudarė *22,3*‰, kai ankstyvesniųjų 15 metų vidurkis siekė *27,2*‰. Visai nežinodami, kiek buvo 1937 mt. mirimų ir negalėdami empiriškai nustatyti natūrinio prieauglio skaičių, bet pasiremdami praeities nustatytais hipotetiniais dydžiais ir kvadratingojo nukrypimo apibrėžtų paklaidų ribomis,

Lentelė 40.

PAVYZDINIO SKAIČIAVIMO SCHEMA NATŪRINIO PRIEAUGLIO HIPOTETINIAMS DYDŽIAMS.

Nr.	Metai	Gimimų skirtumai nuo vidurkio (iš lent. 37, sk. 2): $\triangle_x = x_i - \bar{x}$	Natūrinio prieauglio hipotetiniai dydžiai y′: $\bar{y} + \frac{r\sigma_y}{\sigma_x}\triangle_{\bar{x}} = a + bx = y'$ $11{,}6 + 0{,}5\triangle_x = y'$	Natūrinis prieauglis iš stat. stebėjimo (lent. 37, sk. 4) (y_i)	Statistinio ir hipotetinio nat. prieauglio skirtumai: ($y' - y_i$)	Skirtumų kvadratai: $(y' - y_i)^2$
	0	1	2	3	4	5
1	1922	0,1	11,6 + 0,5 . 0,1 = 11,7	9,6	2,1	4,41
2	1923	1,2	11,6 + 0,5 . 1,2 = 12,2	13,3	— 1,1	1,21
3	1924	2,2	11,6 + 0,5 . 2,2 = 12,7	13,1	— 0,4	0,16
4	1925	1,7	11,6 + 0,5 . 1,7 = 12,5	12,0	0,5	0,25
5	1926	1,3	11,6 + 0,5 . 1,3 = 12,3	13,1	— 0,8	0,64
6	1927	2,2	11,6 + 0,5 . 2,2 = 12,7	12,1	0,6	0,36
7	1928	1,6	11,6 + 0,5 . 1,6 = 12,4	13,2	— 0,8	0,64
8	1929	0	11,6 + 0,5 . 0,0 = 11,6	10,1	1,5	2,25
9	1930	0,2	11,6 + 0,5 . 0,2 = 11,7	11,5	0,2	0,04
10	1931	— 0,4	11,6 — 0,5 . 0,4 = 11,4	11,0	0,4	0,16
11	1932	0,1	11,6 + 0,5 . 0,1 = 11,7	12,0	— 0,3	0,09
12	1933	— 1,5	11,6 — 0,5 . 1,5 = 10,8	12,1	— 1,3	1,69
13	1934	— 2,4	11,6 — 0,5 . 2,4 = 10,4	10,2	0,2	0,04
14	1935	— 3,8	11,6 — 0,5 . 3,8 = 9,7	9,4	0,3	0,09
15	1936	— 3,0	11,6 — 0,5 . 3,0 = 10,1	10,6	— 0,5	0,25
			173,9	173,3		12,28

galime šitaip teigti: nemažiau kaip 66% tikimybės yra užtai, kad 1937 mt. natūrinis prieauglis sudarė apie 9,15‰, ir apie 99,7% tikimybės yra už tai, kad natūrinio prieauglio tikrasis koeficientas nuo mūsų spėjamojo (9,15‰) skirtumas nesudarė daugiau kaip 23‰. Patikrinkime:

1. 1922—1936 mt. gimimų lyginio vidurkis sudarė 27,2‰, o 1937 mt. teišėjo tik 22,3%, taigi skirtumas $\triangle_x = x = -4{,}9‰$.

2. To paties periodo natūrinio prieauglio y vidurkis $\bar{y} = 11{,}6‰$, ir regresijos formuloj natūrinio prieauglio hipotetiniams dydžiams surasti konstanta b = 0,5.

3. Tuo būdu 1937 mt. natūrinio prieauglio hipotetinis dydis turi išeiti:

$$y = \bar{y} + b\,x = 11{,}6 + 0{,}5\,(-4{,}9) = 11{,}6 - 0{,}5 \cdot 4{,}9 =$$
$$= 11{,}6 - 2{,}45 = 9{,}15.$$

4. Hipotetinių dydžių vidurinis kvadratingasis nukrypimas $\sigma\,(=)\,0{,}9$. Iš to seka mūsų spėjamo natūrinio prieauglio tikimybės ir jos paklaidų šios ribos (pagal trijų sigmų taisyklę):

0,68 arba 68% tikimybė, kad 1937 mt. natūrinis prieauglis buvo $9{,}15 \pm 1 \cdot 0{,}9 = 8{,}25 - 10{,}05‰$;

0,95 arba 95% tikimybė, kad 1937 mt. natūrinis prieauglis buvo $9{,}15 \pm 2 \cdot 0{,}9 = 7{,}25 - 10{,}95‰$;

0,997 arba 99,7% tikimybė, kad 1937 mt. natūrinis prieauglis buvo $9{,}15 \pm 3 \cdot 0{,}9 = 6{,}45 - 11{,}85‰$.

5. Patikrinę gi statistikos šaltinius randame, kad 1937 m. natūrinis prieauglis siekė 23 113 asmenys ir tai sudarė 9,148%, t. y. beveik lygiai tiek pat, kaip anksčiau išskaičiuotas hipotetinis dydis (paklaida nesiekia nė 1‰).

Aplamai, toks nepaprastas hipotetinių dydžių sutapimas yra retas atsitikimas; ir galimas jis dažniau tik tais atvejais, kai hipotetiniai dydžiai skaičiuojami perdaug nenutolus nuo bazinio periodo. Tuo pačiu aprėžiamos ir hipotetinių dydžių pritaikymo ribos.

Taip pat įsidėmėtina, kad st. eilutės hipotetinių dydžių skirtumai yra mažesni, kaip skirtumai iš originalių skaičių eilutės,

o bendroji hipotetinių dydžių suma atitinka originalių skaičių sumą (pasitaiką skirtumai, kaip ir lentelės 2 ir 3 skiltyse, tegali pasireikšti tik artutinio skaičiavimo neišvengiamų paklaidų ribose). Dėl šios priežasties koreliacijos koeficientas gali būti panaudotas ir eilučių linkmės kitimų atsitiktiniams nukrypimams išlyginti.

P. S. Literatūra ir papildomosios žinios randamos aplamai ten pat, kaip ir praeitame skyriuje apie koreliacijas (p. 319).

D. Statistikos eilučių išlyginimai ir pagrindinės linkmės apibrėžimai.

13. St. eilučių išlyginimų esmė ir statistinė prasmė.

Statistikos eilučių išlyginamasis procesas vyksta ir kiekybinėj klasifikacijoj. Kai originaliniai dydžio vienetai suskirstomi vienodais intervalais į paruošiamąsias dydžio grupes, pirmykštės eilutės dydžiai jau išeina daugiau ar mažiau „išlyginti", t. y. st. stebėjimo metu gautųjų originalių dydžių mažesnieji skirtumai paneigiami ir mažiau ar daugiau eliminuojami, užtat juo ryškiau pabrėžiamas stambesniais skirtumais dydžio grupių lyginamasis svoris.

Tačiau šios rūšies statistikos eilučių išlyginimas yra ne tam tikros statistikos eilutės lyginamojo metodo tikslas ir ne statistinio analizio uždavinys, o tik kito, arba kitais tikslais daromo, statistinio uždavinio sprendimu gautasis jau antraeilis padarinys. Tam tikrais atvejais šis antraeilis statistinio analizio rezultatas gali būti panaudotas ir specifiniam statistikos eilučių išlyginimu spręstinam uždaviniui. Pav., yra žinoma, kad visuotinio gyventojų surašymo keliu gautosios metinės amžiaus grupės, kurių intervalai žymimi 0, visuomet būna žymiai didesnės, kaip gretimosios amžiaus grupės, kurių intervalai atžymėti skaitmeniu 9 arba 1. Šis reiškinys pareina nuo to, kad beveik visi žmonės lengviau prisimena apytikrius dydžius, ir greičiau yra linkę sakyti, kad X ar Y turi 30 ar 50 metų, kaip 29 arba 31 ar 49 arba 51 metų. Dėl šios priežasties beveik visuomet originaliniai amžiaus duomenys rodo nepaprastai didelį gausingumą visų metinių amžiaus grupių, ku-

rių intervalas aprėžtas 0 (10, 20, 30 ir t. t.) ir nepaprastai mažos (skystos) būna joms artimiausios grupės (su skaičiais 9 ir 11, 19 ir 21 ir t. t.). Aišku, kad nepaprastas nulinių (0) grupių prasikišimas tegalimas tik kaimyninių grupių sąskaiton. Tokiu atveju sakoma, kad statistinis stebėjimas yra davęs sistematinio pobūdžio paklaidą, kuri yra reikalinga pašalinti. Pašalinti gi arba „ištaisyti" šios rūšies paklaidas galima ne tik specialiai pritaikintu išlyginimo metodu, bet ir klasifikuojant į amžiaus klases nemažesniais kaip 10 metų intervalais. Nes jeigu, pav., 20—30 amžiaus klasėj 20-nėj amžiaus grupė bus prasikišusi ne tik dėl to, kad jon pateko nemaža vienetų iš 21-nės amžiaus grupės, bet ir iš 19-nės, tat bendrasis tos klasės gyventojų skaičius bus išlygintas vėl silpniau atstovaujamos 29-nės amžiaus grupės, kurios dalis bus patekusi jau į 30-nę amžiaus grupę, priklausančią 30—40 metų amžiaus klasės.

Vis dėlto šios rūšies išlyginimu galima tenkintis tik tam tikrais atsitikimais, be to, kai pasistatytam uždaviniui pakanka tik apytikrių duomenų. Nes lengva pastebėti, kad sistematinio pobūdžio paklaidos, pasikartojančios kiekvienoj dešimtoj metinio amžiaus grupėj, 10-metinėse amžiaus klasėse tik tuo atveju gali būti pakankamai išlygintos, kai 0-inių dydžio grupių prasikišimai visoj statistinėj eilutėj yra vienodi, t. y. nulemti a b i e j ų gretimųjų grupių. Bet jeigu, pav., 19 metų amžiaus grupė būtų atidavusi 20 metų grupei tik 1000, o 29 metų amžiaus grupė 30-ies metų amžiaus grupei — 3 000 vienetų, tat amžiaus klasė nuo 20 ligi 30 metų jau būtų dirbtinai sumažinta, o gretimoji amžiaus klasė — padidinta. Iš šio pavyzdžio matome, kad st. eilučių išlyginimas (arba išsilyginimas) vienodų intervalų dydžio grupėmis gali būti nepakankamas.

S t a t i s t i k o s e i l u č i ų i š l y g i n i m o t i k s l a s b e i u ž d a v i n y s gali būti ne vienas ir nevienodas. Visų pirma, jei turimuose statistikos duomenyse yra atsitiktinio ar sistematinio pobūdžio skirtingais intervalais pasikartojančių paklaidų, — reikalingi išlyginimo metodai, kurie šias paklaidas pašalintų. Bet gali būti ir taip, kad turimuose statistikos duomenyse st. stebėjimo nei klasifikacijos sudarytų paklaidų visai nėra, o vis dėlto jų „išlyginimas" neišvengiamas, jei norima kitimų pagrindinė linkmė griežčiau apibrėžti arba iš originalinės

st. eilutės eliminuoti periodiškai pasikartojančius nukrypimus arba vad. sezoninius svyravimus, kad paryškėtų charakteringieji, pav., konjunktūrinės arba struktūrinės kilmės kitimai.

Taigi st. eilučių išlyginimo statistiniai uždaviniai gali būti įvairūs arba skirtingi. Dėlto yra reikalingi ir skirtingi st. eilučių išlyginimo metodai. Statistiškai metodologiniu atžvilgiu st. eilučių išlyginimo metodus tenka į dvi pagrindines grupes suskirstyti: a) vieną grupę sudaro metodai st. eilučių pagrindinei linkmei apibrėžti, ir b) antrą — metodai, kurie tik smulkesnius, atsitiktinių arba laikinių aplinkybių (veiksnių) sudarytus nukrypimus eliminuoja.

14. Statistikos eilučių pagrindinės linkmės apibrėžimas.

Paimkime mūsų jau analizuotus Lietuvos natūrinio gyventojų judėjimo duomenis 1922—1936 mt. laikotarpyje. Prisižiūrėję natūrinio prieauglio sk. eilutei mes matome, kad 1922 mt. natūrinis prieauglis tesudarė tik 9,6‰, sekančiais metais staiga pakilo ligi 13,3‰, ir po to vėl, pamažu krisdamas, sugrįžo ligi 10,1‰; paskui vėl pakilo ligi 12,1‰, ir po to vėl ėmė kristi. Kokia gi bendroji natūrinio prieauglio per šį laikotarpį tendencija, t. y. visų kitimų pagrindinė linkmė? Dar neaiškesnė jungtuvių st. eilutės kitimų tendencija. O tuo tarpu šitą tendenciją, t. y. kitimų pagrindinę linkmę, kiek galint tiksliau nustatyti bei griežčiau apibrėžti yra būtinai reikalinga, jei norima tinkamai ir teisingai natūrinio gyventojų judėjimo tendencijas įvertinti, ar bet kurias socialkultūrinės politikos priemones tikslingai panaudoti.

Kai norima tik bet kurios rūšies kitimų pagrindinę linkmę tiksliau apibrėžti, statistinio analizio metodas privalo kiek galint ideališkiau eliminuoti visus — tiek atsitiktinių, tiek antraeilių veiksnių sudarytus — „nuo normos“, šiuo atveju nuo pagrindinės linkmės, nukrypimus. Tokie nukrypimai gali pareiti: a) nuo statistinio stebėjimo paklaidų, b) nuo statistinės klasifikacijos ar analizio metodų netinkamumo, c) nuo klasifikuojamo skaičiavimo paklaidų ir d) nuo mažiau ar daugiau periodiškai pasikartojančių pašalinių arba vad. antraeilių veiksnių įtakos. Pav., pas mus gimimų statistika yra susieta su naujagimių krikštu; tat gali įsibrauti (ir įsibrauja) gimimų statistikoj stebė-

jimo paklaidų tiek, kiek naujagimių krikštijama ne tuojau po gimimo, o ilgesniam laikui praslinkus. Statistinės klasifikacijos bei analizio metodų natūrinio gyventojų judėjimo statistikoj paklaidų gali būti tiek, kiek, pav., neatsižvelgiama mechaninio gyventojų judėjimo kitimų natūrinio gyventojų judėjimo lyginius arba koeficientus skaičiuojant. Periodiškai pasikartojančių veiksnių įtaka mūsų natūrinio gyventojų judėjimo statistikoj ypač ryški mėnesiniuose duomenyse, kai dėl tam tikrų priežasčių lapkričio ir ypač vasario mėnesiais kas metai būna labai daug, o gruodžio ir kovo-balandžio — nepaprastai mažai vedybų.

Pastarosios kilmės periodiškai pasikartojantieji nuo normos nukrypimai mūsų natūrinio judėjimo statistikoj savaime išsilygina, kai duomenys patiekiami ne atskiriems mėnesiams, o ištisiems metams. Tačiau periodiškai pasikartojančių veiksnių įtakos gali būti ir metiniuose duomenyse (pav. dėl konjunktūros svyravimų įtakos vedyboms bei gimimams). O jei paimsime ne natūrinio gyventojų judėjimo st. duomenis, o iš ūkio reiškinių statistikos, tai ilgesniais ar trumpesniais periodais pasikartojančių konjunktūros veiksnių įtaka bus visai aiški ir neginčijama. Dėl šios priežasties yra reikalingas pagrindinės linkmės kitimams apibrėžti toksai metodas, kuris funkcionuotų bet kuriais atvejais vienodai ir patikimai. Šiuos visus reikalavimus patenkina vad. mažiausiųjų kvadratų metodas.

a) Mažiausiųjų kvadratų metodas pagrindinei linkmei pabrėžti.

Jau patsai metodo apibrėžimas rodo, kad šis metodas siejamas su „kvadratais", ir ne su bet kokiais, o su „mažiausiais" kvadratais. Matematinis šio metodo pagrindimas, kaip ir kiti matematinio analizio metodai statistikoj, glūdi tikimybių bei vad. diferencialiame skaičiavime. Šioj vietoj pakaks pabrėžti, kad pagrindinė ir statistikai pati svarbioji šio metodo ypatybė yra ta, kad mažiausiųjų kvadratų metodu išlygintos skaičių eilutės kitimų pagrindinė linkmė nuo originalinės sk. eilutės kitimų linkmės būna nutolusi visų mažiausia, ir tuo būdu gautieji skirtumai iš originalinės ir išlygintosios skaičių eilu-

tės taip pat būna mažiausi ta prasme, kad joks kitas išlyginimo metodas mažesnių skirtumų duoti negali.

Kad būtų aiškesnė šio metodo esmė ir statistinė prasmė, paimkime jau minėtus natūrinio gyventojų prieauglio duomenis y_i, kur 15 metų aritmetinis vidurkis $\bar{y} = 11{,}6‰$, o kraštutiniai nukrypimai siekė 9,4 — 13,3 promilių. Iš statistinių vidurkių bei st. eilučių dispersijos teorijos mes žinome, kad ideališkiausias bet kurios st. eilutės reprezentantas yra aritmetinis vidurkis ir kad iš viso mažiausi st. eilutės dydžių nukrypimai — tai skirtumai nuo aritmetinio vidurkio. Šitą tiesą žinant prašyte prašosi mintis, kad ir st. eilučių pagrindinei linkmei apibrėžti geriausia turėtų tikti skirtumai nuo vidurkio.

Savaime ši mintis nėra neteisinga. Tačiau, jeigu mes norėtume šiais skirtumais operuoti kitimų pagrindinės linkmės ieškodami, tai, visų pirma, sutiktume nenugalimas techninio pobūdžio kliūtis. Nes, kaip žinoma, algebrinė skirtumų nuo vidurkio suma, jei tik skaičiavimo paklaidos nepadaryta, visuomet lygi 0. Nes kai $\bar{x} = \frac{\Sigma x_i}{n}$, tai

$$\Sigma(x_i - \bar{x}) = (x_1 - \bar{x}) + (x_2 - \bar{x}) + \ldots +$$
$$+ (x_i - \bar{x}) + \ldots + (x_n - \bar{x}) = 0.$$

Jei gi imsime nuo vidurkio skirtumus paneigdami ± reiškinius, panašiai kaip rodo formula XXIII a aritmetinių nukrypimų sumai $\Sigma | x_i - \bar{x} |$ arba jų vidurkiui ($v = \frac{\Sigma| x_i - \bar{x} |}{n}$, tai tie skirtumai nepasakys, kuria kryptimi st. eilutės kitimai vyksta. Be to, jei imti vidutinį aritmetinį nukrypimą v ir padaryti jį linkmės kitimų tiesės konstanta b, panašiai kaip tiesės lygtyj ($y = a + bx$), kur $a = \bar{y}$, o x — nuotoliniai dydžiai nuo paties vidurio, tai gausime tiesę, bet n e i š l y g i n t ą skaičių eilutę. Pav., mūsų y-nėj eilutėj, kur $\bar{y} = 11{,}6$,

$$v = \frac{| x_i - \bar{x} |}{n} = \frac{16{,}9}{15} = 1{,}13, \quad n = 15 \text{ ir } x_i \text{ eilutė seka kaip}$$

$x_{-7}, x_{-6}, \ldots x_{-1}, x_0, x_1, \ldots x_7$, kraštutiniai dydžiai gaunami:
$y_7 = 11{,}6 + 1{,}13 \cdot 7 (=) 19{,}5$ ir $y_{-7} = 11{,}6 - 1{,}13 \cdot 7 (=) 3{,}7$.

O tai jau visai neįmanomi skaičiai, nes tokių didelių nukrypimų per visą 15 metų visai nėra buvę.

Formula statistinių eilučių pagrindinei linkmei apibrėžti privalo šių ypatybių: a) kitimų linkmę apibrėžiantieji skaičiai privalo kiek galint mažiau nutolti nuo originalinės st. eilutės skaičių, b) išlygintos bei pagrindinę linkmę vaizduojančios skaičių eilutės suma turi atitikti originalinės skaičių eilutės sumą, c) išlygintos skaičių eilutės skirtumai nuo vidurkio, taip pat ir vidurinis kvadratingasis nukrypimas bei dispersijos koeficientas, turi būti mažesnis kaip tos pačios rūšies dydžiai iš originalinės skaičių eilutės ir d) pati formula ir skaičiavimo technika turi būti kiek galint paprastesnė, nes pagrindinės linkmės apibrėžimas statistikoj priklauso elementarinių ir dėlto kuo plačiausia naudotino st. analizio priemonių grupės.

Šias visas ypatybes mažiausiųjų kvadratų metodas ir turi, o kai kurie šio metodo variantai įgalina net labai ilgas, t. y. labai didelį narių skaičių n turinčias, st. eilutes greitai ir lengvai išlyginti.

Pats pirmasis ir pagrindinis st. eilutės kitimų linkmės apibrėžimas — tai linkmės kitimų tiesė pagal žinomą formulą, kur

$$y = a + bx.$$

b) St. eilučių pirmaeilė linkmė arba Linkmės kitimų tiesė.

St. eilučių pirmaeilė linkmė, arba linkmės kitimų tiesė, gali būti apibrėžiama iš algebros žinomų vad. normaliųjų regresijos lygčių pagalba, kur tiesės lygties reiškiniai

$$y = a + bx$$

gaunami:

$$\Sigma y = na + b\Sigma x$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \qquad \text{(XLV).}$$

Skaičiavimo pavyzdžiui paimkime jau naudotus natūrinio gyventojų prieauglio duomenis, kur $n = 15$, taigi $x_n = x_{15}$. Tuo būdu

$$\Sigma y = 15a + b\Sigma x = 15a + (1 + 2 + 3 + \ldots + 15)\, b =$$
$$= 15a + 120b,$$

ir

$$\Sigma xy = 120a + b\Sigma x^2$$
$$= 120a + b\,(1^2 + 2^2 + 3^2 + \ldots + 15^2)$$
$$= 120a + 1240b.$$

Linkmės kitimų tiesės skaičiavimo schema pavaizduota lentelėje 41.

Lentelė 41.

NATŪRINIO PRIEAUGLIO LINKMĖS KITIMŲ TIESĖS SKAIČIAVIMO SCHEMA NORMALIŲ REGRESIJOS LYGČIŲ PAGALBA.

Eilės Nr.	Metai	Natūrinis prieauglis (y)	Nat. prieauglio eilutės seka (x)	Sandauga (xy)	Eil. sekos nuotolinių skaičių kvadratai (x^2)	Linkmės kitimų tiesė $a + bx = y'$ $12{,}6 - 0{,}132x\ (=)\ y'$
n	0	1	2	3	4	5
1	1922	9,6	1	9,6	1	12,6—0,132.1 (=) 12,47
2	1923	13,3	2	26,6	4	12,6—0,132.2 (=) 12,34
3	1924	13,1	3	39,3	9	 12,20
4	1925	12,0	4	48,0	16	 12,07
5	1926	13,1	5	65,5	25	 11,94
6	1927	12,1	6	72,6	36	 11,81
7	1928	13,2	7	92,4	49	 11,68
8	1929	10,1	8	80,8	64	 11,54
9	1930	11,5	9	103,5	81	 11,41
10	1931	11,0	10	110,0	100	 11,28
11	1932	12,0	11	132,0	121	 11,15
12	1933	12,1	12	145,2	144	 11,02
13	1934	10,2	13	132,6	169	 10,88
14	1935	9,4	14	131,6	196	 10,75
15	1936	10,6	15	159,0	225	12,6—0,132.15 = 10,62
		173,3= $=\Sigma y$	120= $=\Sigma x$	1 348,7= $=\Sigma xy$	1 240= $=\Sigma x^2$	173,16

Iš lentelės išrašę atitinkamus skaičius gauname:

$$\Sigma y = na + b\Sigma x = 15a + 120b = 173{,}2;$$
$$\Sigma xy = a\Sigma x + b\Sigma x^2 = 120a + 1240b = 1348{,}7.$$

Kad gauti reikalingus skaičius nežinomiems reiškiniams a ir b, visų pirma tenka nuveikti šias skaičiavimo operacijas:

a) rasti x-inės eilutės dydžių vidurkį $\bar{x}$,

b) gautuoju vidurkiu sudauginti pirmosios lygties visus skaičius,

c) iš b operacijos gautųjų lygties skaičių atimti antrosios lygties skaičius, kad gauti b reiškinį atstovaujančius skaičius.

Šį uždavinį spręsdami gauname:

a) $\bar{x} = \frac{\Sigma x}{n} = \frac{120}{15} = 8;$

b) $\bar{x}na + \bar{x}(b\Sigma x) = \bar{x}\Sigma y$

$= 8 . 15a + 8 . 120b = 8 . 173{,}2$

$= 120 + 960b = 1385{,}6;$

c) $120a + 1240b = 1348{,}7$

$-120a + 960b = 1385{,}6$

$280b = -36{,}9$

Kai žinome, jog $280b = -36{,}9$, tai viena b

$$b = \frac{-36{,}9}{280} (=) -0{,}132.$$

O jeigu $b = -0{,}132$ ir, kaip anksčiau matėme,

$$120a + 1240b = 1348{,}7,$$

tai iš to naunama:

$$\begin{aligned} 120a + 1240\,(-0{,}132) &= 1348{,}7 \\ 120a - 163{,}64 &= 1348{,}7 \\ 120a &= 1348{,}7 + 163{,}64 \\ 120a &= 1512{,}34 \end{aligned}$$

Taigi a:

$$a = \frac{1512{,}34}{120} (=) 12{,}6.$$

Tuo būdu kitimų linkmei tiesės lygtis gaunama pavidalo:

$$y = a + bx = 12{,}6 - 0{,}132x.$$

Linkmės kitimų tiesė pasikliaujant normalios regresijos lygtimis turi tą gerą ypatybę, kad pačios lygtys ir jų sprendi-

mas jau pažįstamas iš vidurinės mokyklos algebros kurso. Tačiau patsai normalinių lygčių sprendimas reikalingas nemaža laiko, — ir juo daugiau, kuo st. eilutės narių skaičius n didesnis. Be to, kai n didelis, tenka operuoti labai dideliais skaičiais. Šiuo atžvilgiu tiesės formulai reikalingieji reiškiniai daug paprasčiau gaunami paimant y-inės eilutės aritmetinį vidurkį dėl a ir trupmeną iš xy sumų sandaugos, padalintos x-inės eilutės kvadratų suma, dėl b. Tuo būdu iš karto gauname:

$$a = \bar{y} \quad \text{ir} \quad b = \frac{\Sigma xy}{\Sigma x^2}.$$

Visa tiesės lygtis išeina:

$$y = a + bx = \bar{y} + \frac{\Sigma xy}{\Sigma x^2}\, x \qquad \text{XLVa).}$$

X-inės eilutės nuotoliniai dydžiai rašomi su neigiamais ar teigiamais ženkleliais po lygų skaičių į abi puses pradedant nuo st. eilutės paties vidurio:

$$x_{-n},\ x_{-(n-1)},\ \ldots,\ x_{-2},\ x_{-2}\ x_0,\ x_1,\ x_2,\ \ldots,\ x_{n-1}$$

Schematinis skaičiavimo pavyzdys patiektas lentelėj 42, p. 335.

Kai turime duomenis reiškiniams:

$$\Sigma y = 173{,}3, \quad \Sigma xy = -37{,}7 \quad \text{ir} \quad \Sigma x^2 = 280,$$

tai linkmės kitimų tiesei gauti reikalingieji skaičiai konstantoms a ir b gauname.

$$a = \bar{y} = \frac{\Sigma y}{n} = \frac{173{,}3}{15} (=)\ 11{,}553$$

$$b = \frac{\Sigma xy}{\Sigma x^2} = \frac{-37{,}7}{280} (=) -0{,}135.$$

Tuo būdu tiesės lygtis išeina:

$$y = a + bx = \bar{y} + \frac{\Sigma xy}{\Sigma x^2} x$$

$$= 11{,}553 - 0{,}135x.$$

Pagal šią lygtį natūrinio prieauglio linkmės kitimų tiesės y dydžiai, surašyti lentelės 42 skiltyj 5, išeina beveik tokie pat,

Lentelė 42.

LINKMĖS KITIMŲ TIESĖS SKAIČIAVIMO SCHEMA MAŽIAUSIŲJŲ KVADRATŲ METODU, KAI N YRA NELYGINIS SKAIČIUS.

Eilės Nr.	Metai	Natūrinis prieauglis (y)	Nuotolis nuo eilutės vidurio (x)	Sandauga (yx)	Nuotolinių dydžių kvadratai (x^2)	Linkmės kitimų tiesė a+bx=y 11,553-0,135x=y′
n	0	1	2	3	4	5
1	1922	9,6	—7	—67,2	49	12,50
2	1923	13,3	—6	—79,8	36	12,36
3	1924	13,1	—5	—65,5	25	12,23
4	1925	12,0	—4	—48,0	16	12,09
5	1926	13,1	—3	—39,3	9	11,96
6	1927	12,1	—2	—24,2	4	11,82
7	1928	13,2	—1	—13,2	1	11,69
8	1929	10,1	0	0	0	11,55
9	1930	11,5	1	11,5	1	11,42
10	1931	11,0	2	22,0	4	11,28
11	1932	12,0	3	36,0	9	11,15
12	1933	12,1	4	48,4	16	11,01
13	1934	10,2	5	51,0	25	10,88
14	1935	9,4	6	56,4	36	10,74
15	1936	10,6	7	74,2	49	10,61
		$173,3=\Sigma y_i$		$-37,7=\Sigma xy$	$280=\Sigma x^2$	173,29

kaip lentelėj 41 (maži skirtumai susidaro dėl artutinio skaičiavimo duomenų trumpinimo).

Iš patiektosios lentelės 41, kur linkmės kitimų tiesė išskaičiuota pagal formulą XLVa, aiškiai matyti šios formulos geros ypatybės: a) didesnis pavaizdumas ($a = \bar{y}$ ir konstantą b atstovauja betarpiškai iš dviejų skilčių dalmuo), b) mažiau darbo, c) operuojama mažesniais skaičiais. Šio metodo silpnoji pusė tai ta, kad x-inės eilutės nuotolinių dydžių žymėjimo sistemą tenka pakeisti, jeigu y-inės eilutės narių skaičius n yra lyginis (porinis). Tada x-iniai nuotoliai rašoma:

$$x_{-n},\ x_{-(n-2)},\ \ldots,\ x_{-1},\ x_1,\ x_3,\ \ldots,\ x_n.$$

Priežastis ta, kad lyginės skaičių eilutės y viduriniai dydžiai y_{-1} ir y_1, kai nėra ir negali būti tarpinio nario y_0, skiriasi 2 b, kai kiti gretimieji dydžiai (pav., y_{-1}, y_{-2}... arba y_1, y_2 ir t. t.) tesiskiria tik 1 b. Dėlto porinėj eilutėj reikalinga ir

kitų y_i narių nuotoliams x paimti tokią pat distanciją (per du vienetu).

Lyginės eilutės nuotolinių dydžių narius galima žymėti ẋ. Tada $\dot{x} = 2x$ ir tuo būdu kitimų tiesės formulą bus galima rašyti šitaip:

$$y = a + b\dot{x}$$

Lentelė 43.

LINKMĖS KITIMŲ TIESĖS SKAIČIAVIMO SCHEMA MAŽIAUSIŲJŲ KVADRATŲ METODU, KAI N YRA LYGINIS SKAIČIUS.

Eilės Nr.	Metai	Natūrinis prieauglis (y)	Nuotol. nuo eil. vidurio ẋ	Sandauga (yẋ)	Nuotolių kvadr. ($\dot{x}^2$)	Linkmės kitimų tiesė: a + bx = y 11,4—0,0825ẋ = y′
	0	1	2	3	4	5
1	1922	9,6	— 15	— 144,0	225	12,64
2	1923	13,3	— 13	— 172,9	169	12,47
3	1924	13,1	— 11	— 144,1	121	12,31
4	1925	12,0	— 9	— 108,0	81	12,14
5	1926	13,1	— 7	— 91,7	49	11,98
6	1927	12,1	— 5	— 60,5	25	11,81
7	1928	13,2	— 3	— 39,6	9	11,65
8	1929	10,1	— 1	— 10,1	1	11,48
9	1930	11,5	1	11,5	1	11,32
10	1931	11,0	3	33,0	9	11,15
11	1932	12,0	5	60,0	25	10,99
12	1933	12,1	7	84,7	49	10,82
13	1934	10,2	9	91,8	81	10,66
14	1935	9,4	11	103,4	121	10,49
15	1936	10,6	13	137,8	169	10,33
16	1937	9,1	15	136,5	225	10,16
		182,4		— 112,2	1 360	182,40

Linkmės kitimų teisę lyginio skaičiaus n eilutėj, pratęsę natūrinio prieauglio duomenis ligi 1937 mt., gauname:

$$a = \bar{y} = \frac{182,4}{16} = 11,4,$$

$$b = \frac{\Sigma \dot{x}y}{\dot{x}^2} = \frac{-112{,}2}{1360} = -0{,}0825.$$

Tuo būdu tiesės lygtis,

$$y = a + b\dot{x},$$

pakeitus simbolius skaičiais, išeina:

$$y = 11{,}4 - 0{,}0825\,\dot{x}.$$

Skaičiavimo rezultatai matyti iš lentelės 43, sk. 5 (pusl. 336).

c) Statistikos eilučių antraeilė, trečiaeilė ir t. t. linkmė.

St. eilučių pirmaeilė linkmė, arba linkmės kitimų tiesė, parodo tik pagrindinę kitimų krypties tendenciją. Tačiau gali būti reikalinga pažinti ne tik pagrindinę linkmę, vaizduojamą linkmės kitimų tiesės, bet ir charakteringus tokių kitimų vingius, arba žymesnius nuo pagrindinės linkmės nukrypimus. Pav., stebint natūrinio prieauglio duomenis tuojau į akį krinta ta aplinkybė, kad 1922 mt. nat. prieauglis tesudarė tik 9,6‰, paskui tuojau pakilo ligi 13‰ su viršum, po to kiek krito ir vėl pakilo ligi 13,2‰ 1928 m., bet 1935 m. vėl nukrito ligi 1922 m. lyginio. Taigi čia aiškiai turime ne nuolatinę vienos linkmės kitimų tendenciją, o dvi: 1922—1928 mt. ir 1928—1936 mt. Norėdami abiejų periodų tendencijas atvaizduoti išlygintoj st. eilutėj, turime pasikliauti jau ne tiesės lygtimi, o aukštesnės eilės parabolės lygtimis.

Aplamai, st. eilutėms išlyginti ir bet kurios eilės kitimų linkmei apibrėžti galioja šios formulos:

$$y_0 = a_0 = \bar{y}$$

$$y_1 = a + bx$$

$$y_2 = a + bx + cx^2$$

$$y_3 = a + bx + cx^2 + dx^3$$

. . . . . .

. . . . . .

. . . . . .

$$y_n = a + bx + cx^2 + \ldots + a_n x^n.$$

Sulig tuo, ar st. eilutė išlyginama jos pačios aritmetiniu vidurkiu ($y_0 = \bar{y}$), tiese (y_1) ar bet kurios eilės parabole, galima vadinti st. eilutės nuline (y_0), pirmaeile (y_1), antraeile (y_2), trečiaeile (y_3) ar bet kurios kitos eilės (y_n) linkme ligi tol, kol apibrėžtoji n-tosios linkmės kreivė paseks net originalinių st. eilutės dydžių atsitiktinius nukrypimus.

Šios rūšies st. eilučių išlyginimo bei kitimų linkmės apibrėžimo metodai gali būti pagrindžiami baigtinių skirtumų bei mažiausiųjų kvadratų skaičiavimo principais. Taip pat ne tik pirmaeilė, bet ir antraeilė ar bet kurio kito laipsnio st. eilučių linkmė gali būti nustatoma normalių regresijos lygčių pagalba. Tačiau iš anksčiau patiektų pavyzdžių lengva įsitikinti, kad normalių regresijos lygčių pagalba net pirmaeilės linkmės išskaičiavimas jau gana komplikuotas ir nemaža atima darbo. Be to, statistikos praktikoj, ypač kai st. eilutės narių skaičius yra didelis ir iš anksto negalima žinoti, kurios eilės linkmė yra statistikai prasmingiausia, visuomet tenka mažiau ar daugiau eksperimentuoti. Tuo tarpu vad. normalios lygtys, davusios, sakysim, antraeilei linkmei reikalingas konstantas a, b ir c, padaro jas (ir visą padėtą darbą) bevertėmis tų pačių konstantų trečiaeilei arba ketvirtaeilei linkmei. Šiais ir dar kitais sumetimais normalios lygtys st. eilučių linkmei apibrėžti gana sunkiai pritaikomos ir dėlto st. praktikoj jos beveik nevartojamos.

d) St. eilučių išlyginimas bei pagrindinės linkmės apibrėžimas kintamaisiais vidurkiais.

Pritaikomosios statistikos tose srityse, kur st. vaizdavimas yra reikalingas mažiau ar daugiau išlygintų st. eilučių (be atsitiktinių ar sistematinių nukrypimų), dažniausia panaudojama kintamųjų vidurkių viena ar kita sistema. Ligi šiol statistikos praktikoj plačiausia vartojama kintamųjų vidurkių sistema galima apibendrinti šitokia formula:

$$y'_i = \frac{1}{N}(y_{i-n} + \ldots + y_{i-1} + y_i + y_{i+1} + \ldots + y_{i+n}), \quad \text{(XLVI)}.$$

Šioj formuloj N atstovauja kintamajam vidurkiui naudojamųjų narių skaičių, o n — skaičius narių, kurie yra kairėj (—n) arba dešinėj (+n) nuo „lyginamojo" nario y_i. Sulig tuo, ar norima st. eilutės kitimų linkmės dydžius daugiau ar mažiau priartinti originaliems st. eilutės dydžiams, kintamiems vidurkiams imamųjų narių skaičius n turi būti didesnis ar mažesnis. Pav., jei norima gauti linkmės st. eilutę, kur būtų eliminuoti tik atsitiktiniai nukrypimai, pakanka kintamųjų vidurkių iš trijų, daugiausia penkių, st. eilutės narių. Bet jeigu norima eliminuoti ir mažiau ar daugiau sistematinio pobūdžio nukrypimus (pav., sezoninius ar kt. rūšių periodinius svyravimus), kintamiesiems vidurkiams tenka imti atitinkamai didesnį st. eilutės narių skaičių. Pagrindinis dėsnis, kurio tenka tokiais atvejais atsižvelgti, yra šis: kintamiesiems vidurkiams imamųjų st. eilutės narių skaičius N turi būti nuo 1½ ligi 2 kartų didesnis už eliminuotinų nukrypimų pasikartojimus. Pav., jeigu kurios nors rūšies nukrypimai (sezoniniai svyravimai) pasikartoja kas 3—4 n (mėnesis), tai kintamiesiems vidurkiams reikalingas minimalinis narių skaičius $N \geqq 4 + \frac{4}{2} = 6$. Jeigu nukrypimai sistematiškai pasikartoja maždaug kas šeštasis sk. periodas, tai minimalinis $N \geqq 6 + \frac{6}{2} = 9$, ir t. t. Dėl to, pav., ir konjunktūros statistikoj, kur dažniausia kintamaisiais vidurkiais eliminuojami sezoniniai, maždaug kas 5—8 mėn. pasikartojantieji nukrypimai, kintamieji vidurkiai skaičiuojami iš 12 st. eilutės dydžių. Tačiau, jeigu N yra lyginis skaičius, tai XLVI formulą tenka modifikuoti šiuo būdu:

$$y'_i = \frac{1}{N}\left(\frac{1}{2}y_{i-n} + \ldots + y_{i-1} + y_i + y_{i+1} + \ldots + \frac{1}{2}y_{i+n}\right). \qquad \text{(XLVIa).}$$

Jeigu, tarsime, kad y-inės eilutės turi narių, kurie išrykiuotoj eilutėj atstovaujami $y_1, y_2, \ldots, y_{25}$, ir norėsime išlyginti trijų dydžių kintamaisiais vidurkiais, tai N = 3. Tuo būdu XLVI formula gaus šitokį konkretų pavidalą:

$$\acute{y}_2 = \frac{1}{3}(y_{2-1} + y_2 + y_{2+1}) = \frac{1}{3}(y_1 + y_2 + y_3),$$

$$\acute{y}_3 = \frac{1}{3}(y_{3-1} + y_3 + y_{3+1}) = \frac{1}{3}(y_2 + y_3 + y_4),$$

.....

.....

.....

$$\acute{y}_{24} = \frac{1}{3}(y_{24-1} + y_{24} + y_{24+1}) = \frac{1}{3}(y_{23} + y_{24} + y_{25}).$$

Jei gi norėtume tą pačią eilutę išlyginti keturių dydžių kintamaisiais vidurkiais, tai išlyginimo formulą tektų pakeisti šitaip:

$$\acute{y}_3 = \frac{1}{4}\left(\frac{1}{2}y_{3-2} + y_{3-1} + y_3 + y_{3+1} + \frac{1}{2}y_{3+2}\right) =$$

$$= \frac{1}{4}\left(\frac{1}{2}y_1 + y_2 + y_3 + y_4 + \frac{1}{2}y_5\right);$$

$$\acute{y}_4 = \frac{1}{4}\left(\frac{1}{2}y_2 + y_3 + y_4 + y_5 + \frac{1}{2}y_5\right)$$

.....

.....

.....

$$\acute{y}_{23} = \frac{1}{4}\left(\frac{1}{2}y_{21} + y_{22} + y_{23} + y_{24} + \frac{1}{2}y_{25}\right).$$

Iš formulų išvystymo matyti, kad patys kraštutiniai y-inės eilutės dydžiai palieka neišlyginti. Ir juo didesnis N, tuo didesnis kraštutinių dydžių skaičius negali būti išlygintas. Pav., jeigu $N = 2$ arba 3, tai — neišlyginti palieka y_0 ir y_n, o jeigu $N = 4$ arba 5, tai jau nelyginami y_0, y_1, iš vieno st. eilutės galo ir y_{n-1} su y_n iš antrojo galo (du pirmutiniai ir du paskutiniai dydžiai). O jeigu $N = 12$ arba 13, tai palieka neišlyginti pirmieji šeši ir paskutiniai šeši dydžiai. Dėl šios priežasties kintamųjų vidurkių metodas st. praktikoj naudotinas tik tais atvejais, kai pirmųjų ir paskutinių st. eilutės dydžių išlyginimas nesudaro st. analizio ar vaizdavimo būtinos sąlygos.

Kintamaisiais vidurkiais st. eilučių išlyginimas gali būti daromas ir vad. s v e r t i n i a i s k i n t a m a i s i a i s v i d u r-

k i a i s — tais atvejais, kai išlyginamojo nario (dydžio) lyginamasis svoris norima tik „aplyginti", bet ne visai išlyginti. Statistinės prasmės svertiniai kintamieji vidurkiai daugiau turi tik atsitiktinius nukrypimus eliminuojant, t. y. tik kelių — st. praktikoj, paprastai, tik trijų — dydžių kintamaisiais vidurkiais. Tokiu atveju „aplyginamas" dydis padvigubinamas, ir apibendrinta kintamųjų vidurkių formula gauna šį varijantą:

$$\acute{y}_i = \frac{1}{N}(y_{i-n} + \ldots + y_{i-1} + 2y_i + y_{i+1} + \ldots + y_{i+n}) \quad \text{(XLVIb)}.$$

Kai statistikos praktikoj, paprastai, kintamieji svertiniai vidurkiai skaičiuojami tik iš trijų st. eilutės narių, tai ši specifinė svertinių kintamųjų vidurkių formula gauna šitokį pavidalą:

$$\acute{y}_i = \frac{1}{4}(y_{i-1} + 2y_i + y_{i+1}) \; (=) \frac{1}{2}\left(\frac{1}{2}y_{i-1} + y_i + \frac{1}{2}y_{i+1}\right).$$

Kintamaisiais vidurkiais išlyginimo metodui pavaizduoti galime paimti tuos pačius natūrinio gyventojų prieauglio duomenis, kurie buvo naudojami ir linkmės kitimų tiesei apibrėžti (žiūr. lent. 44 pusl. 342).

15. Kintamųjų parametrų metodas st. eilutėms išlyginti.

Rusų jaunųjų statistikų grupė 1936 mt. savo išleistoj „Statistikoj" demonstruoja eilučių išlyginimo metodą, kuris pasižymi savo paprastumu ir gautųjų rezultatų patenkinamumu. Rusai, berods, šį metodą demonstruoja ne st. eilučių linkmės charakteringiems kitimams pabrėžti, o st. dydžių dispersijai intervaluose paryškinti bei išlyginti. Tačiau, turint galvoje, kad st. eilučių intervalų dispersiją išlyginamieji metodai savo esme nesiskiria nuo st. eilučių išlyginimo metodų ir paprastai skaičiavimas atliekamas pasikliaujant tomis pačiomis, ar kiek modifikuotomis, formulomis, tenka tą metodą patikrinti. Pirmiausia pravartu prisižiūrėti ir patikrinti jo pirmykštį pavartojimą originalioje statistikos lentelėje 45 (p. 342)[14].

[14] Žiūr. Statistika, psl. 301—303.

Lentelė 44.

ST. EILUČIŲ IŠLYGINIMO SCHEMA KINTAMAISIAIS VIDURKIAIS.

Nr.	Metai	Natūrinis prieauglis (‰) (y_i)	Išlyginimas trijų dydžių kintamaisiais vidurkiais $\acute{y}_i=\frac{1}{3}(y_{i-1}+y_i+y_{i+1})$	Išlyginimas keturių dydžių kintamaisiais vidurkiais $\acute{y}_i=\frac{1}{4}(\frac{1}{2}y_{i-2}+y_{i-1}+y_i+y_{i+1}+\frac{1}{2}y_{i+2})$
0	1	2	3	4
1	1922	9,6	—	—
2	1923	13,3	12,0	—
3	1924	13,1	12,8	12,44
4	1925	12,0	12,7	12,70
5	1926	13,1	12,4	12,59
6	1927	12,1	12,8	12,36
7	1928	13,2	11,8	11,93
8	1929	10,1	11,6	11,59
9	1930	11,5	10,9	11,30
10	1931	11,0	11,5	11,40
11	1932	12,0	11,7	11,49
12	1933	12,1	11,4	11,13
13	1934	10,2	10,6	10,75
14	1935	9,4	10,1	10,20
15	1936	10,6	9,7	—
16	1937	9,1	—	—
		182,4	162,0	139,88

Lentelė 45.

KREIVĖS ORDINATŲ DISPERSIJOS APIBRĖŽIMAS TREJETUKO KINTAMAISIAIS INTERVALAIS.

Intervalai	Intervalo eil. nr. i	Skaičius w_i	$\frac{1}{24}2(w_{i-1}-2w_i+w_{i+1})$	Ordinata y_i
0—1	1	59	$-\frac{1}{24}(0-118+162)=-2$	57
1—2	2	162	$-\frac{1}{24}(59-324+237)=+1$	163
2—3	3	237	$-\frac{1}{24}(162-474+211)=+4$	241
3—4	4	211	$-\frac{1}{24}(237-422+140)=+2$	213
4—5	5	140	$-\frac{1}{24}(211-280+85)=-1$	139
5—6	6	85	$-\frac{1}{24}(140-170+48)=-1$	84
6—7	7	48	$-\frac{1}{24}(85-96+24)=-1$	47
7—8	8	24	$-\frac{1}{24}(48-48+14)=-1$	23
8—9	9	14	$-\frac{1}{24}(24-28+7)=0$	14
9—10	10	7	$-\frac{1}{24}(14-14+6)=0$	7
10—11	11	6		

Metodo veikimas patiektoj lentelėj 45 matyti aiškiausia: vieni originalios eilutės dydžiai truputį sumažinami, kiti kiek padidinami, treti — palieka nepaliesti. Aplamai, tačiau, skirtumas išlygintųjų ir neišlygintųjų dydžių yra visai nedidelis. Bet šis, palyginti, mažas skirtumas yra, kaip lengva įsitikinti, ne metodo charakteringa ypatybė, o pareina vien tik nuo skaičiavimo pagrindan paimtos konstantos $=\frac{1}{24}$. Tačiau ši konstanta nieko bendra neturi nei su bendruoju dydžio grupių skaičiumi, kurių yra 11, nei su koreguojamiems dydžiams imamų st. eilutės narių skaičiumi (narių paimta 3, iš kurių vidurinis padvigubinamas). Taip pat skaičiaus 24 nėra galima logiškai susieti su bet kokia šių ar kitų duotosios st. eilutės skaičių kombinacija (pav., bendrasis vienetų skaičius N = 993). O jeigu skaičiavimo pagrindan paimti konstanta ne iš $\frac{1}{24}$, o, tarsime, iš $\frac{1}{4}$, tai skirtumai tuojau padidės, ir, pav., y_1 išeis ne 57, o 48:

$$y_1 = 59 - \frac{1}{4}(0 - 118 + 162) = 59 - 11 = 48.$$

Taip pat y_3 išeis ne 241, o

$$y_3 = 237 - \frac{1}{4}(162 - 474 + 211) = 237 - 25 = 262.$$

Ir t. t.

Iš čia matome, kad rusų statistikų gautieji maži skirtumai tėra savivalingas, be jokio tvirto kriterijaus, apibrėžto argumento konstantos padarinys. Savaime tasai rusų statistikų pademonstruotas metodas jų patiekto pavyzdinio skaičiavimo uždaviniui spręsti gali būti ir patenkinamas. Bet jo apibendrinimas ir konstantos dogmatinis piršimas be jokio metodologinio ar matematiškai teoretinio pagrindimo daro tą metodą jo pirmykštėj formoj statistikos eilučių išlyginimo praktikai, ypač st. eilučių linkmės apibrėžimui, nepriimtiną. Juo labiau, kad st. eilučių bet kurio laipsnio linkmei apibrėžti formula, kaip anksčiau buvo parodyta, be kitų ypatybių, yra reikalinga sąlygos, kad linkmės variantų skirtumų bendroji suma išeitų kiek galint mažesnė, — ir jau būtinai mažesnė už skirtumų sumą iš

originalios st. eilutės dydžių. Tuo tarpu rusų demonstruotasis metodas mažesniuosius dydžius sumažino, o pačius didžiuosius — dar daugiau padidino.

Bet ši siūlomo metodo veikimo ypatybė lengvai pašalinama, nes ji pareina vien tik nuo to, kad formulos antroji dalis padaryta priklausoma nuo konstantos $\frac{1}{24}$ su neigiamu pažymiu. Pačių rusų patiektoji formula, ją apibendrinus, atrodo šitaip (simboliai iš originalo):

$$y_i = w_i - \frac{1}{24}(w_{i-1} - 2w_i + w_{i+1}).$$

Tačiau, jeigu lygties dešinėjėj pusėj argumento konstantos neigiamą pažymį (—) pakeisime teigiamu (+) ir parašysime:

$$y_i = w_i + \frac{1}{24}(w_{i-1} - 2w_i + w_{i+1}),$$

tai formulos veikimo rezultatai griežtai pasikeis: neigiami skaičiai pasidarys teigiamais, o teigiami — neigiamais. Efektas gi bus tas, kad st. eilutės sekoj didesnieji dydžiai bus mažinami, o mažesnieji — didinami. O tai ir yra pagrindinė st. eilučių išlyginimo sąlyga.

Tuo būdu rusų statistikų kolektyvo demonstruotasis metodas st. eilučių išlyginimo ir linkmės apibrėžimo reikalui duodasi lengvai pritaikomas. Palieka, tačiau, metodologiškai neleistinas beprincipiškumas bei savivalingumas parenkant argumentus konstantai. Nes jeigu patiektoji konstanta $\frac{1}{24}$ duotosios st. eilutės skaičiams išlyginti ir būtų galima laikyti visai tinkama, tai, paėmus tą pačią konstantą st. eilutei išlyginti ir tuo atveju, kai skaičiai būtų, sakysim, vidutiniškai 10 kartų didesni ar mažesni, — skirtumai padidėtų arba sumažėtų. Dėlto ir patiektoji išlyginimo formula savo statistinės prasmės nustotų.

Bet ir ši formulos neigiamybė gali būti pašalinta. Prisižiūrėjus rusų statistikų kolektyvo demonstruotojo metodo veikimui lengva pastebėti, kad tasai metodas savo esmėje labai primena kintamųjų vidurkių išlyginimo metodą, — ypač vad.

svertinių kintamųjų vidurkių metodą, kai patsai lyginamasis dydis imamas dvigubai didesnis (žiūr. form. XLVIb).

Kadangi aptariamuoju atveju išlyginimui imami tik trys st. eilutės dydžiai, iš kurių išlyginamasis atsveriamas dvigubai ir tuo būdu $N = 3 + 1 = 4$, tai svertinių kintamųjų vidurkių formula rašytina šitaip:

$$\acute{y}_i = \frac{1}{4}(y_{i-1} + 2y_i + y_{i+1}).$$

Jei pamsime tą patį kriterijų konstantai ir rusų statistikų kolektyvo formulai, tai ji, pačių rusų rašyboj, atrodys šitaip:

$$y_i = w_i + \frac{1}{4}(w_{i-1} - 2w_i + w_{i+1}).$$

Tačiau rusų naudojamą simbolį w, kai jų formulą taikome y-inei eilutei išlyginti, galima pakeisti y, o tuo pačiu visa formula dar labiau prisiartina kintamųjų vidurkių formulai. Nes tikrenybėje:

$$\acute{y}_i = \frac{1}{4}(y_{i-1} + 2y_i + y_{i+1}) = y_i + \frac{1}{4}(y_{i-1} - 2y_i + y_{i+1}) =$$

$$= w_i + \frac{1}{4}(w_{i-1} - 2w_i + w_{i+1}),$$

t. y. tapatybės.

Lengva įsitikinti, skaičiavimo pagrindan paėmus tuos pačius duomenis.

Jeigu $y_1 = 59$, $y_2 = 162$, $y_3 = 237$, $y_4 = 211$ ir t. t., kaip parodyta lentelėj 45, tai atitinkami $\acute{y}_i$ iš kintamųjų svertinių vidurkių pagal formulą:

$$\acute{y}_i = \frac{1}{4}(y_{i-1} + 2y_i + y_{i+1})$$

išeina:

$$\acute{y}_1 = \frac{1}{4}(y_0 + 2y_1 + y_2) = \frac{1}{4}(0 + 2 \cdot 59 + 162) = \frac{280}{4} = 70,$$

$$\acute{y}_2 = \frac{1}{4}(59 + 2 \cdot 162 + 237) = \frac{620}{4} = 155,$$

$$\acute{y}_3 = \frac{1}{4}(162 + 2 \cdot 237 + 211) = \frac{847}{4} = 211{,}75$$

Ir t. t.

Tiems patiems skaičiams rusiškoji formula:

$$\acute{y}_i = y_i + \frac{1}{4}(y_{i-1} - 2y_i + y_{i+1})$$

duoda:

$$\acute{y}_1 = y_1 + \frac{1}{4}(y_{1-1} - 2y_1 + y_2) = 59 + \frac{1}{4}(0 - 2 \cdot 59 + 162) =$$

$$= 59 + \frac{1}{4}(-118 + 162) = 59 + \frac{44}{4} = 70,$$

$$\acute{y}_2 = 162 + \frac{1}{4}(59 - 2 \cdot 162 + 237) = 162 + \frac{1}{4}(-28) =$$

$$= 162 - 7 = 155,$$

$$\acute{y}_3 = 237 + \frac{1}{4}(162 - 2 \cdot 237 + 211) = 237 - 101 : 4 = 211{,}75$$

Ir t. t.

Tuo būdu įsitikiname, kad rusų statistikų kolektyvo siūlomas metodas st. eilutėms išlyginti iš tikrųjų tebėra tik daugiau komplikuotas specifinio tipo svertinių kintamųjų vidurkių metodo variantas, — jei tą metodą statistiškai matodologiniu atžvilgiu įprasminti per argumento konstantą, parinkus ją tuo pačiu kriterijum, kaip parenkama trijų dydžių svertiniams kintamiems vidurkiams gauti.

Bet šiuo atveju specifinis abiejų metodų panašumas ir baigiasi. Nes jeigu paimti paprastus trijų dydžių kintamuosius vidurkius $\acute{y}_i = \frac{1}{3}(y_{i-1} + y_i + y_{i+1})$ ir tiems dydžiams išskaičiuoti $\acute{y}_i$ pagal atitinkamai pritaikintą trijų kintamųjų parametrų metodo formulą, tai skaičiavimo rezultatai nesutaps. Pav.:

$$\acute{y}_i = \frac{1}{3}(y_1 + y_2 + y_3) = \frac{59 + 162 + 237}{3} (=) 153, \quad \text{o}$$

$$y'_i = y_2 + \frac{1}{3}(y_1 - y_2 + y_3) = 59 + \frac{59 - 162 + 237}{3} = 193.$$

Aplamai imant, rusų statistikų kolektyvo demonstruotas metodas, — jį modifikavus bei eliminavus statistiškai metodologiniu atžvilgiu nepateisinamą ir neleistiną savivalingumo elementą, — naudotinas statistinėms eilutėms išlyginti bei kitimų linkmei apibrėžti visais tais atvejais, kai tiesės linkmė dėl vienų ar kitų priežasčių laikoma nepakankama, o norima eliminuoti ne vien atsitiktinius nukrypimus siauresniąja prasme, bet ir pastovesnių veiksnių įtaką neperilgame laikotarpyje.

16. Sumavimų metodas statistikos eilutėms išlyginti bei kitimų linkmei aprėžti.

Jau buvo minėta, kad matematiškai tobulas statistinių eilučių išlyginimas ne tiesės linkme tėra galimas pasikliaujant parabolės lygtimis. Bet aukštesniųjų eilių parabolės lygtimis statistikos eilučių lyginimas reikalingas ne tik daug darbo, bet yra ir gana komplikuotas dėl vis naujų, pereinant į aukštesnę eilę, konstantų įvedimo ir kt. Be to, statistikos praktikoj nuolatos tenka susidurti su reikalu išlyginamos statistinės eilutės narių skaičių vis plėsti ir plėsti sulig naujai gaunamais statistinio stebėjimo rezultatais (pav., visas vad. konjunktūros statistinis stebėjimas bei vaizdavimas). Ir šitokių nuolat ilgėjančių eilučių išlyginimas aukštesnės eilės parabolės lygčių pagalba statistikui yra nepatogus ir ištaigingas dėl darbo gausumo ir jo komplikuotumo.

Šios rūšies ydų mažiausia turi vadinamas sumavimų metodas, kuris tuo pačiu laiku atitinka visus vad. st. matematiškojo analizio metodui statomus reikalavimus.

Sumavimų metodo teoretinis pagrindimas ir X_i funkcijos, taip k_i ir l_i faktorių, matematinis išvystymas gana populiariai yra išdėstytas L o r e n z'o veikale „T r e n d“, kur yra patiektos ir atitinkančios Četverikovo schemos, kontrolinės formulos, taip pat ir X_i funkcijos bei k_i ir 1_i faktorių lentelės. Ten pat yra išdėstyta, kuo būdu šitos sumavimų metodų gaunamos linkmės tiesės bei kreivės interpoliuojamos, koreliuoja-

mos ir t. t. Statistikos pagrinduose visa tai atpasakoti arba pakartoti būtų tiek pat netikslinga, kaip būtų netikslinga atpasakoti matematikos ar algebros pagrindimo dėsnius arba pakartoti bei išnagrinėti vad. normalias lygtis, logaritmų ir kt. matematinių dydžių sudarymo formulas ar p. Dėl šios priežasties šioj vietoj teks pasitenkinti tik pagrindinių sumavimo metodo ypatybių apibrėžimu tiek, kiek tos ypatybės gali turėti reikšmės statistikos eilučių išlyginimui įprasminti.

Pirmutinė ir pati svarbioji šio metodo ypatybė yra ta, kad čia vartojamieji funkciniai bei faktoriniai dydžiai pareina tik nuo statistinės eilutės narių bendrojo skaičiaus *(N)*, visai neatsižvelgiant jos atskirų variantų n_i dydžio nei lyginamojo svorio. Iš to seka, kad vieną kartą išskaičiuotas bet kuriam statistinės eilutės narių N skaičiui bet kurios eilės funkcinis dydis X_i galioja kiekvienos statistinės eilutės tam pačiam narių skaičiui N, visai nepriklausomai nuo tos eilutės atskirų dydžių lyginamojo svorio, nei nuo jų pasikeitimų eilutės sekoj.

Sumavimų metodo pranašumas statistikos eilutėms išlyginti iš dalies jau matyti iš lygčių sugretinimo. Pav., pasikliaujant normaliomis lygtimis norint pereiti iš pirmaeilės (tiesės) y_1 į antraeilę linkmę y_2 (parabolę), tenka iš naujo skaičiuoti konstantas a ir b, be to, įvesti dar trečią konstantą — c. Tuo tarpu sumavimų metode tiesės linkmei (y_1) išskaičiuotoji konstanta a_1 galioja ir antraeilei parabolei (y_2), o tereikalinga tik konstanta a_2. Tas pats dėsnis galioja pereinant į trečiaeilę parabolę y_3 ir t. t.

Aplamai, bet kurios eilės linkmei skaičiuoti galioja šitokios lygtys:

$$y_0 = \bar{y} = a_0$$

$$y_1 = a_1 + b_1x = a_0 + a_1X_1 = y_0 + a_1X_1$$

$$y_2 = a_2 + b_2x + c_2x^2 = a_0 + a_1X_1 + a_2X_2 = y_1 + a_2X_2$$

$$y_3 = a_3 + b_3x + c_3x^2 + d_3x^3 = a_0 + a_1X_1 + a_2X_2 + a_3X_3 = y_2 + a_3X_3$$

$$y_4 = a_4 + b_4x + c_4x^2 + d_4x^3 + e_4x^4 = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 = y_3 + a_4X_4$$

Ir t. t.

Aplamai, sumavimų metodo lygtis apibendrinant galima rašyti:

$$y'_i = y'_{i-1} + a_i X_i \qquad \text{(XLVIII)}.$$

Iš lygčių suderinimo jau pakankamai aiškiai matyti sumavimų metodo pranašumas, — ypač jei prisiminsime, kad funkciniai dydžiai X_i gali būti tiesiog paimti iš atitinkamų lentelių. Pagaliau, X_i funkcija, reikale, ir išskaičiuoti nesunku. Pav., bet kuriam st. eilutės narių n e l y g i n i a m s k a i č i u i X_i funkcijai gauti galioja šios formulos:

$$X_1 = x\sqrt{\frac{12}{N^2-1}}, \quad X_2 = -\sqrt{\frac{5\,(N^2-1)}{16\,(N^2-4)}},$$

$$X_3 = -\sqrt{\frac{7\,(N^2-1)\,(N^2-9)}{256\,(N^2-4)}}$$

Ir t. t.

Lyginiams skaičiams galioja panašios konstrukcijos formulos, kurios jau savo išviršine forma bei pasikartojančiomis konstantomis rodo artimos giminystės ryšius su binomu bei tikimybėmis. Tuo ir pagrindžiamas sumavimų metodo matematinis apytikrumas, o tuo pačiu — ir jų statistinis patikimumas.

Aplamai, čia patiektos X_i funkcinių dydžių formulos tėra tik kontrolinės. Jos duoda pagrindą duotosios eilutės narių skaičių (N) atitinkančius $x_{-n}, \ldots, x_{-1}, x_0, x_1, \ldots, x_n$ funkcinius dydžius X_i nustatyti (jeigu N yra lyginis skaičius, tai X_i ir x_i dydžių eilutės seka: $x_{-n}, \ldots, x_{-3}, x_{-1}, x_1, x_3, \ldots, x_n$). Patys X_i funkcijos dydžiai ligi X_5 dėl N ligi 25 patiekiami atskirai iš minėtos Lorenz'o knygos, kurioj yra skaičiai ligi X_6 net dėl N = 80.

Patsai sumavimų metodu skaičiavimas pradedamas nustatymu skaičių S_i reiškiniams. Sulig tuo, kuriuo laipsniu statistinę eilutę norima išlyginti, reikalinga ir S_i. S_0, kuri yra reikalinga konstantai a_0 arba vidurkiui, atstovaujama visų statistinės eilutės n dydžių bendrosios sumos. Jei statistinės eilutės narius ar jų grupes pažymėsime y_i, tat $S_0 = \Sigma y_i$. Skaičiui, atstovaujančiam reiškinį S_1, yra reikalinga

padaryti papildomą sumavimą sulig x-iniais nuotoliais nuo vidurinio dydžio, panašiai kaip sumavimų metodu kvadratingojo nukrypimo ieškant. Panašiu būdu gaunami skaičiai reiškiniams S_2, S_3 ir t. t. Skirtumas tas, kad lyginiams S_i (S_2, S_4 ir t. t.) reikalingieji skaičiai gaunami teigiamų ir neigiamų s_i sudėtimi, o nelyginiams (S_1, S_3 ir t. t.) — atimtimi.

Aplamai, kumuliuotoms sumoms S_i sudaryti galioja šios lygtys:

$S_0 = s_0 + s_{-0}$ (arba $s_0 + y_0 + s_{-0}$, jeigu N nelyginis sk.)

$S_1 = s_1 - s_{-1}$, $S_2 = s_2 + s_{-2}$, $S_3 = s_3 - s_{-3}$,

$S_4 = s_4 + s_{-4}$ ir t. t.

o aplamai

$$S_i = s_i \pm s_{-i} \qquad \text{(XLIX)}.$$

Turint kumuliuotas sumas S_i, lengvai randamos reikalingos konstantos a_i. Pav., linkmės tiesei gauti galioja formula XLVIII:

$$y_1 = a_0 + a_1 X_1.$$

Kai žinoma, kad

$$a_0 = \bar{y} = \frac{\Sigma y_i}{N} = \frac{S_0}{N},$$

tai reikalinga tik konstanta a_1 ir funkcinis skaičius X_1. Jeigu N yra n e l y g i n i s, tai konstanta a_1 gaunama pagal formulą:

$$a_1 = \frac{1}{N}\sqrt{\frac{12}{N^2-1}} \cdot S_1 \qquad \text{(La)}.$$

Jei N lyginis, tai ta pati konstanta a_1 gaunama šitaip:

$$a_1 = \frac{1}{N}\sqrt{\frac{3}{N^2-1}}\,(2S_1 - S_0) \qquad \text{(Lb)}.$$

Funkciniai dydžiai X_i nelyginiams skaičiams gaunami pagal formulą:

$$X_1 = x\sqrt{\frac{12}{N^2-1}},$$

ir lyginiams N

$$X_1 = x \sqrt{\frac{3}{N^2 - 1}}, \qquad \text{(LIa ir b)}$$

kur x yra nuotoliniai eilutės skaičiai.

Iš pirmosios eilės kumuliuotos sumos S_1 gauti konstantai a_1 pagalbinė formula

$$\frac{1}{N}\sqrt{\frac{12}{N^2 - 1}} \left(\text{arba } \frac{1}{N}\sqrt{\frac{3}{N^2 - 1}} \right)$$

atstovauja faktorių k_1, kurio visi reikalingieji variantai randami atitinkamose lentelėse. Statistikos praktikoje, kad nereikėtų operuoti smulkiomis trupmenomis, paprastai įvedamas skaičius 1000. Tuo būdu

$$k_1 = \sqrt{\frac{12}{N^2 - 1}} \cdot \frac{1000}{N} \text{ arba } \sqrt{\frac{3}{N^2 - 1}} \cdot \frac{1000}{N}.$$

Tada, tačiau, norint gauti a_1 tenka imti ne S_1, o $\frac{S_1}{1000}$

Panašiu būdu gaunamos ir konstantos a_2, a_3 ir kt. Tik aukštesnės eilės funkciniai dydžiai X_i kiekvienam nuotoliniam skaičiui galioja betarpiškai. Pav.:

$$X_2 = - \sqrt{\frac{5\,(N^2 - 1)}{16\,(N^2 - 4)}},$$

t. y. jau be x, kuris galioja daugikliu atitinkamai formulai dėl X_1.

Kad būtų aiškesnė sumavimų metodo esmė ir statistinė prasmė, paimkime vėl tuos pačius natūrinio prieauglio duomenis nuo 1922 ligi 1937 mt.

Statistinės eilutės narių skaičius šiuo atveju yra $15 = N$, t. y. nelyginis. Iš ankstyvesnių skaičiavimų jau žinome, kad bendroji 15 metų natūrinio prieauglio lyginių suma $\Sigma y_i = 173{,}3$.

Tad:

$$S_0 = 173{,}3, \quad \text{o} \quad a_0 = \bar{y} = \frac{173{,}3}{15} = 11{,}55.$$

Konstantai a_1 gauti reikalinga S_1, o $S_1 = s_1 - s_{-1}$. Kai N yra nelyginis, tai $S_1 = s_1 - s_{-1} = \Sigma xy$, mums pažįstama

iš mažiausiųjų kvadratų metodo tiesės linkmei skaičiuoti. O iš 334 psl. mes žinome, kad $\Sigma xy = -37{,}7$, tad

$$S_1 = -37{,}7.$$

Bet mums reikalinga a_1, kuriai panaudojama formula La psl. 350. Kai N = 15, tai

$$a_1 = \frac{1}{N}\sqrt{\frac{12}{N^2-1}} \cdot S_1 = \frac{1}{15}\sqrt{\frac{12}{15^2-1}}(-37{,}7) =$$

$$(=)\frac{1}{15}\sqrt{\frac{12}{224}}(-37{,}7)\ (=)\ \frac{-37{,}7 \cdot 0{,}2315}{15}\ (=) -0{,}582.$$

Tuo tarpu funkcinis dydis X_1, kai N = 15, gaunamas:

$$X_1 = x\sqrt{\frac{12}{15^2-1}} = x\sqrt{\frac{12}{224}}\ (=)\ x\ \frac{3{,}46}{14{,}97}\ (=)\ 0{,}2315x.$$

Tuo būdu tiesės linkmei lygtis išeina:

$$y_1 = a_0 + a_1X_1\ (=)\ 11{,}55 - 0{,}582 \cdot 0{,}2315x$$
$$(=)\ 11{,}55 - 0{,}135x,$$

t. y., galų gale, gauti tie patys dydžiai kaip mažiausiųjų kvadratų metodu (žiūr. psl. 334).

Sumavimų metodo veikimas ligi trečiaeilės parabolinės linkmės y_3 nelyginiam N yra demonstruotas „Socialekonominėj Statistikoj" 516—521 p. p. Papildomai čia patiekiu sumavimų metodo veikimą natūriniam gyventojų judėjimui 1922—1937 mt., kur N = 16, t. y. N yra lyginis skaičius ir dėlto skaičiavimo technika šiek tiek skiriasi. Visi skaičiavimai padaryti pagal Čebyševo schemą, o X_i funkcijos ir k_i bei l_i faktoriniai dydžiai panaudoti iš Lorenz'o lentelių.

Kai N = 16, o st. eilutę norime išlyginti ligi ketvirtosios eilės y' imtinai, tai a_i konstantoms reikalingi faktoriniai dydžiai k_i ligi k_4 imtinai, be to, faktoriai l_3 ir l_4, nes:

$$a_0 = \frac{S_0}{N}$$
$$a_1 = R_1$$
$$a_2 = R_2 - R_0$$
$$a_3 = (R_3 - R_1)l_3$$
$$a_4 = (R_4 - 10R_2 + 3R_0)l_4.$$

Lentelė 46.

SUMAVIMŲ METODO SCHEMA NATŪRINIO GYVENTOJŲ PRIEAUGLIO DUOMENIMIS 1922—37

Nr.	Metai	yi	$S_0=s_0+s_{-0}$ $=85,9+96,5$ $=182,4=S_0$	$S_1=s_1—s_{-1}$ $=372,3—337,2$ $=35,1=S_1$	$S_2=s_2+s_{-2}$ $=840,0+992,7$ $=1832,7=S_2$	$S_3=s_3—s_{-3}$ $=2071,0—$ $—1449,6=$ $=621,4=S_3$	$S_4=s_4+s_{-4}$ $=2443,1+$ $+2856,7=$ $=5299,8=S_4$
0	1	2	3	4	5	6	7
1	1922	9,6	9,6	9,6	9,6	9,6	9,6
2	1923	13,3	22,9	32,5	42,1	51,7	61,3
3	1924	13,1	36,0	68,5	110,6	162,3	223,6
4	1925	12,0	48,0	116,5	227,1	389,4	613,0
5	1926	13,1	61,1	177,6	404,7	794,1	1407,1
6	1927	12,1	73,2	250,8	655,5	$s_{-3}=1449,6$	$s_{-4}=2856,7$
7	1928	13,2	86,4	$s_{-1}=337,2$	$s_{-2}=992,7$	—	—
8	1929	10,1	$s_{-0}=96,5$	—	—	—	—
9	1930	11,5	$s_0=85,9$	$s_1=372,3$	—	—	—
10	1931	11,0	74,4	286,4	$s_2=840,0$	$s_3=2071,0$	—
11	1932	12,0	63,4	212,0	553,6	1231,0	$s_4=2443,1$
12	1933	12,1	51,4	148,6	341,6	677,4	1212,1
13	1934	10,2	39,3	97,2	193,0	335,8	534,7
14	1935	9,4	29,1	57,9	95,8	142,8	198,9
15	1936	10,6	19,7	28,8	37,9	47,0	56,1
16	1937	9,1	9,1	9,1	9,1	9,1	9,1

Taigi reikalingi reiškiniai R_j, kurie gaunami iš kumuliuotų sumų, padaugintų faktoriniu dydžiu k_j, šiuo būdu:

$$R_0 = \frac{S_0 k_0}{N}$$

$$R_1 = \frac{(2S_1 - S_0)k_1}{1000}$$

$$R_2 = \frac{S_2 k_2}{1000}$$

$$R_3 = \frac{(2S_3 - S_0)k_3}{1000}$$

$$R_4 = \frac{S_4 k_4}{1000}$$

Faktoriniai dydžiai k_j dėl $N = 16$, išskaičiuoti arba tiesiog iš lentelės išrašyti, yra šie:

$k_0 = 1{,}11144$, $k_1 = 6{,}77908$, $k_2 = 6{,}61570$,
$k_3 = 1{,}07604$, $k_4 = 15{,}4366$.

Taip pat a_3 ir a_4 konstantoms reikalingieji faktoriniai dydžiai l_j gaunami šie:

$l_3 = 2{,}31436$ ir $l_4 = 0{,}330623$.

Turėdami kumuliuotas sumas s_j (žiūr. lent. 46, pusl. 353) pirmiausia gauname S_j ir a_j konstantai reikalingus reiškinius R_j šiuo būdu:

$$S_0 = s_0 + s_{-0} = 85{,}9 + 96{,}5 = 182{,}4$$

$$R_0 = \frac{S_0 k_0}{N} = \frac{182{,}4 \,.\, 1{,}11144}{16} = \frac{202{,}726656}{16} \,(=)\, 12{,}670$$

$$S_1 = s_1 - s_{-1} = 372{,}3 - 337{,}2 = 35{,}1$$

$$R_1 = \frac{2S_1 - S_0}{1000} k_1 = \frac{(2 \,.\, 35{,}1 - 182{,}4) \,.\, 6{,}77908}{1000} =$$

$$= \frac{-112{,}2 \,.\, 6{,}77908}{1000} \,(=)\, -0{,}761$$

$$S_2 = s_2 + s_{-2} = 840{,}0 + 992{,}7 = 1832{,}7$$

$$R_2 = \frac{S_2 k_2}{1000} = \frac{1832{,}7 \,.\, 6{,}61570}{1000} (=) \; 12{,}125$$

$$S_3 = s_3 - s_{-3} = 2071{,}0 - 1449{,}6 = 621{,}4$$

$$R_3 = \frac{2S_3 - S_2}{1000} \; k_3 = \frac{(2 \,.\, 621{,}4 - 1832{,}7) \,.\, 1{,}076}{1000} =$$

$$= \frac{-589{,}9 \,.\, 1{,}076}{1000} = -0{,}635$$

$$S_4 = s_4 + s_{-4} = 2443{,}1 + 2856{,}7 = 5299{,}8$$

$$R_4 = \frac{S_4 k_4}{1000} = \frac{5299{,}8 \,.\, 15{,}4366}{1000} (=) \; 81{,}811$$

Turėdami sumas S_i ir skaičius reiškiniams R_i gauname konstantas a_i šiuo būdu:

$$a_0 = \frac{S_0}{N} = \frac{182{,}4}{16} = 11{,}4$$

$$a_1 = R_1 (=) - 0{,}761$$

$$a_2 = R_2 - R_0 = 12{,}125 - 12{,}670 = -0{,}545$$

$$a_3 = (R_3 - R_1) 1_3 = (-0{,}635 - 0{,}761) \,.\, 2{,}31436 = 0{,}292$$

$$a_4 = (R_4 - 10R_2 + 3R_0) \,.\, 1_4 = (81{,}811 - 10 \,.\, 12{,}125 + 3 \,.\, 12{,}670) \,.\, 0{,}330623 = -0{,}473.$$

Tuo būdu natūrinio gyventojų prieauglio kitimų linkmės, ligi y_4 imtinai, lygtis išeina:

$$y' = a_0 + a_1 X_1 + a_2 X_2 + a_3 X_3 + a_4 X_4$$

$$= 11{,}4 - 0{,}761 \, X_1 - 0{,}545 \, X_2 + 0{,}292 \, X_3 - 0{,}473 \, X_4.$$

Bet iš psl. 349 jau žinome, kad funkcinis dydis X_1 tiesės linkmei galioja tik kartu su nuotoliniais reiškiniais x. Taigi tiesės linkmę skaičiuojant galioja formula:

$$y'_1 = a_0 + a_1 X_1(x) = 11{,}4 - 0{,}761 \, X_1 \, (x).$$

Tuo tarpu kai N = 16, tai

$$X_1 = x \sqrt{\frac{3}{16^2 - 1}} = x \sqrt{\frac{3}{255}} (=) \; x \, \frac{1{,}7321}{15{,}9687} (=) \; 0{,}108465x.$$

Lentelė 47.

SUMAVIMŲ METODU GAUTIEJI TIESINĖS (y_1) IR ANTRAEILĖS (y_2), TREČIAEILĖS (y_3) BEI KETVIRTAEILĖS (y_4) LINKMĖS DUOMENYS.

Nr.	Metai	Nat. prieauglis (‰) y_i	Nuotolis nuo vidurio x	$a_0+a_1X_i(x_i)=y_1$ $=11{,}4-0{,}76$, $0{,}1085x\ (-)11{,}4$ $-0{,}0825\ x=y_1$	X_2	$y_1+a_2X_2=y_2$ $=y_1-0{,}545X_2=$ (y_2)	X_3	$y_3=y_2+a_3X_3$ $=y_2+0{,}292X_3$ (y_3)	X_4	$y_4=y_3+a_4X_4$ $=y_3-0{,}473X_4$ (y_4)
0	1	2	3	4	5	6	7	8	9	10
1	1922	9,6	—15	12,64	1,85	11,62	1,81	11,09	1,59	10,34
2	1923	13,3	—13	12,47	1,11	11,86	0,36	11,75	—0,53	12,00
3	1924	13,1	—11	12,31	0,48	12,05	0,57	12,22	—1,29	12,83
4	1925	12,0	— 9	12,14	0,05	12,17	1,06	12,48	—1,17	13,03
5	1926	13,1	— 7	11,98	—0,48	12,24	1,20	12,59	—0,59	12,87
6	1927	12,1	— 5	11,81	—0,79	12,25	1,06	12,56	0,13	12,50
7	1928	13,2	— 3	11,65	—1,01	12,20	0,71	12,41	0,75	12,06
8	1929	10,1	— 1	11,48	—1,11	12,09	0,25	12,16	1,10	11,64
9	1930	11,5	1	11,32	—1,11	11,93	—0,25	11,86	1,10	11,34
10	1931	11,0	3	11,15	—1,01	11,70	—0,71	11,49	0,75	11,11
11	1932	12,0	5	10,99	—0,79	11,43	—1,06	11,12	0,13	11,06
12	1933	12,1	7	10,82	—0,48	11,08	—1,20	10,73	—0,59	11,00
13	1934	10,2	9	10,66	—0,05	10,69	—1,06	10,38	—1,17	10,92
14	1935	9,4	11	10,49	0,48	10,23	—0,57	10,06	—1,29	10,65
15	1936	10,6	13	10,33	1,11	9,72	0,36	9,83	—0,53	10,07
16	1937	9,1	15	10,16	1,85	9,14	1,81	9,67	1,59	8,94
	—	$182{,}4=\Sigma y_i$	—	$182{,}40=\Sigma y_{i,1}$	—	$182{,}40=\Sigma y_{i,2}$	—	$182{,}40=\Sigma y_{i,3}$	—2	$182{,}36=\Sigma y_{i,4}$

Taigi tiesės linkmei lygtis rašytina šitaip:

$$\acute{y}_1 = 11{,}4 - 0{,}761 \,.\, 0{,}108465\ x\ (=)\ 11{,}4 - 0{,}0825\ x.$$

Tai atitinka tiesės lygtis iš mažiausiųjų kvadratų metodo psl. 336 ir tuo būdu dar kartą įrodoma, kad sumavimų metodas savo esmėje yra tiek pat griežtas ir patikimas, kaip ir mažiausiųjų kvadratų metodas.

Galutinai visa linkmės kitimų lygtis rašytina:

$$y'\ (=)\ 11{,}4 - 0{,}0825\ x - 0{,}545\ X_2 + 0{,}292\ X_3 - 0{,}473\ X_4.$$

Reikalingus dydžius X_2, X_3 ir X_4 dydžius išrašome iš lentelės su pažymiu N = 16, o x_i dydžiai, kai N yra lyginis, eina kaip 1, 3, 5 ir t. t. su neigiamais ir teigiamais reiškiniais lygiai taip pat, kaip ir mažiausiųjų kvadratų metode lentelėje 43 psl. 336. Tenka taip pat turėti galvoje, kad ir aukštesnės eilės X_i funkcijos nelyginiai dydžiai (X_3, X_5 ir t. t.) kinta sulig formula:

$$X_i\ (-x) = -\ X_i(x),$$

kur x yra natūriniai skaičiai.

Tuo būdu iš lentelės išrašytųjų nelyginių funkcijų dydžiai (X_3, X_5 ir t. t.) kairėj pusėj pakeičia reiškinius. Nes jeigu X_3, X_5 ar kt. atžymėta neigiamu reiškiniu, tai, sudauginus kairės pusės nuotolinių x neigiamu reprezentantu x_{-1}, bus gautas teigiamas skaičius, ir, žinoma, išeis neigiamas — jei buvo teigiamas. Dėl to ir lentelės 47 sk. 7. X_3, kai originalus $X_3(-1) = -0{,}251$, pritaikius formulą:

$$X_i\ (-x) = -\ X_i\ (x),$$
$$-0{,}25\ (-1) = 0{,}25.$$

Pav., kai

$$X_{3\,(15)} = 1{,}81,$$

tai

$$X_{3(-15)} = 1{,}81\ (-1) = -\ 1{,}81.$$

Panašiu būdu pasikeičia reiškiniai $X_3(-3)$ ir t. t. Tuo tarpu dešinės pusės X_i funkcijos reiškiniai palieka kaip buvę. Taip pat nesikeičia lyginių funkcijų (X_2, X_4 ir t. t.) reiškiniai, nes čia galioja $X_i\ (-x) = X_i\ (x)$, taigi ir $X_3(-1) = -25 \,.\, 1 =$

$= -0{,}25$, o $X_{31\,(15)} = 1{,}85 \cdot 1 = 1{,}85$. Patys funkciniai dydžiai X_i lentelėje patiekti apytikriai (ligi 0,01 griežtumo). Iš lentelės 47 matyti, kad tik ketvirtaeilės linkmės y_4 dydžių bendroji suma truputį skiriasi nuo originalinės skaičių eilutės (182,36 vietoj 182,40). Šis skirtumas pareina nuo artutinio skaičiavimo, sutrumpinus X_1 dydžius ligi 0,01 griežtumo.

17. Statistinių eilučių linkmės bei išlygintųjų dydžių patikimumo patikrinimas.

Jau buvo minėta, kad st. eilučių išlyginimo svarbiausias uždavinys — eliminuoti atsitiktinių arba ir antraeilės reikšmės nuolatinių, bet mažesniais ar didesniais intervalais pasikartojančių, veiksnių įtaką, kad juo ryškesnė paliktų pagrindinė linkmė arba bendrųjų (konjunktūros ar struktūrinio pobūdžio) veiksnių įtaka. Šios rūšies uždavinys tikslingai išsprendžiamas tegali būti tik tokiais išlyginimo metodais, kurių veikimas pasireiškia kraštutinumus (kraštutinius nukrypimus) išlyginamąja kryptimi. Tad st. eilutėms išlyginti ar kitimų bet kurios eilės pagrindinei linkmei apibrėžti tegali būti naudojami tik tokie metodai, kurių veikimo pasekmėje būtų gaunamos „lygesnės" st. eilutės, negu jos buvo anksčiau (prieš „išlyginimą"). O st. eilučių „lygumo" ir „nelygumo" rodiklis, kaip jau žinome, yra jų dispersijos koeficientas. Vadinasi, „išlygintųjų" st. eilučių dispersijos koeficientas turi išeiti mažesnis, negu neišlygintų, t. y. originalinės skaičių eilutės dispersijos koeficientas.

Patikimiausias dispersijos koeficientas, kaip žinome, yra gaunamas iš vidurinio kvadratingojo nukrypimo. Tačiau šiuo atveju tikslingiau yra ieškoti vid. kvadratingojo nukrypimo ne sulig y'_i dydžių skirtumais nuo kiekvieno eilutės aritmetinio vidurkio, o sulig išlygintos y'_i ir neišlygintos y_i eilučių atitinkamų dydžių savitarpio skirtumais pagal formulą:

$$\delta_{y'} = \sqrt{\frac{\Sigma(y_i - y'_i)^2}{n}} \qquad \text{(LII).}$$

Jei st. eilutė išlyginama keliais metodais, tai šiuo būdu galima ir kiekvieno atskirai paimto išlyginamojo metodo didesnis ar mažesnis tinkamumas bei statistinis prasmingumas nu-

Lentelė 48.

SKAIČIAVIMO SCHEMATINIS PAVYZDYS IŠLYGINTŲJŲ ST. EILUČIŲ SKIRTUMINEI DISPERSIJAI APIBRĖŽTI

Nr.	Metai	Natūrinio prieauglio lyginiai ‰ y_i	$y_i - y_{1,i}$	$(y_i - y'_{1,i})^2$	$y_i - y'_{2,i}$	$(y_i - y'_{2,i})^2$	$y_i - y'_{3,i}$	$(y_i - y'_{3,i})^2$	$y_i - y'_{4,i}$	$(y_i - y'_{4,i})^2$
0	1	2	3	4	5	6	7	8	9	10
1	1922	9,6	—3,04	9,2416	—2,02	4,0804	—1,49	2,2201	—0,74	0,5476
2	1923	13,3	0,83	0,6889	1,44	2,0736	1,55	2,4025	1,30	1,6900
3	1924	13,1	0,79	0,6241	1,05	1,1025	0,88	0,7744	0,27	0,0729
4	1925	12,0	—0,14	0,0196	—0,17	0,0289	—0,48	0,2304	—1,03	1,0609
5	1926	13,1	1,12	1,2544	0,86	0,7396	0,51	0,2601	0,23	0,0529
6	1927	12,1	0,29	0,0841	—0,15	0,0225	—0,46	0,2116	0,40	0,1600
7	1928	13,2	1,55	2,4025	1,00	1,0000	0,79	0,6241	1,14	1,2996
8	1929	10,1	—1,38	1,9044	—1,99	3,9601	—2,05	4,2436	1,54	2,3716
9	1930	11,5	0,18	0,0324	—0,43	0,1849	—0,36	0,1296	0,16	0,0256
10	1931	11,0	—0,15	0,0225	—0,70	0,4900	—0,49	0,2401	—0,11	0,0121
11	1932	12,0	1,01	1,0201	0,57	0,3249	0,88	0,7744	0,94	0,8836
12	1933	12,1	1,28	1,6384	1,02	1,0404	1,37	1,8769	1,10	1,2100
13	1934	10,2	—0,46	0,2116	—0,49	0,2401	—0,18	0,0324	—0,72	0,5184
14	1935	9,4	—1,09	1,1881	—0,83	0,6889	—0,66	0,4356	—1,25	1,5625
15	1936	10,6	0,27	0,0729	0,88	0,7744	0,77	0,5929	0,53	0,2809
16	1937	9,1	—1,06	1,1236	—0,04	0,0016	—0,57	0,3249	0,16	0,0256
		182,4= $=\Sigma y_i$		21,5292 = $=\Sigma(y_i - \acute{y}_1)^2$		16,7528 = $=\Sigma(y_i - \acute{y}_2)^2$		15,3736 = $=\Sigma(y_i - \acute{y}_{3,i})^2$	3,92	11,7742 = $=\Sigma(y_i - \acute{y}_4)^2$

statyti. Pavyzdžiui galime paimti mūsų jau įvairiais metodais analizuotus („lygintus“) natūrinio gyventojų prieauglio duomenis.

Lentelėj 47 yra patiekti išlygintoms natūrinio prieauglio eilutėms skirtumai ir jų kvadratai pradedant tiesės linkme (y_1) ligi ketvirtosios eilės parabolės (y_4). Matome, kad skirtumai, taigi ir viduriniai kvadratingieji nukrypimai (σ), mažėja šiuo būdu:

$$\sigma_1 (=) \sqrt{\frac{21{,}53}{16}} (=) \sqrt{1{,}35} (=) 1{,}16;$$

$$\sigma_2 (=) \sqrt{\frac{16{,}75}{16}} (=) \sqrt{1{,}05} (=) 1{,}02;$$

$$\sigma_3 (=) \sqrt{\frac{15{,}57}{16}} (=) \sqrt{0{,}97} (=) 0{,}98;$$

$$\sigma_4 (=) \sqrt{\frac{11{,}77}{16}} (=) \sqrt{0{,}74} (=) 0{,}86.$$

Tuo būdu dispersijos koeficientas ketvirtaeilės linkmės skaičiams $y_{4,i}$ nesudaro nė 8% (0,86 : 11,4), kai tiesės linkmės gaunama dar daugiau kaip 10% (1,16 : 11,4).

Literatūra ir papildomos žinios: 1) Lorenz, Der Trend, Berlin 1931 (Sonderheft 21 d/ Vierteljahrshefte zur Konjunkturforschung), 2) Rietz-Bauer, Handbuch der mathematischen St. pp. 47—92, 123—158 ir 218—234. Be to, atitinkami skyriai kiekviename moderniškesniame vadovėlyje. Kai kurie autoriai st. eilučių išlyginimą sieja su periodinių bei sezoninių svyravimų eliminavimu. Tačiau šie klausimai, taigi ir literatūra, mano yra nukelti į paskutinį skyrių ir aptariami kartu su sezoninių indeksų sudarymo bei naudojimo klausimais.

E. Statistikos indeksai.

Statistikos indeksai[1] priklauso santykinių skaičių kategorijos — statistikos lyginių (lygmeninių santykio sk.) grupės. Tačiau statistikos indeksai iš kitų santykinių bei lygmeninių skaičių išsiskiria tuo, kad jie, ypač naujesniaisiais laikais, yra paversti mažne universaline statistinio analizio priemone.

[1] Indexziffer, Indexzahlen; nombres indices, indices; Index Numbers.

Jau vien ta aplinkybė, kad yra rimtų ekonomistų, kurie siūlo statistinius indeksus padaryti pinigų vertinamuoju matu, rodo statistinių indeksų ypatingą reikšmę ir svarbą. Dėl to yra mėginimų (Irving Fischer, Flaskämper) net atskirą indeksų teoriją sudaryti.

18. Statistikos indeksų esmė ir rūšys.

Bet kurios rūšies statistinė eilutė, arba tos eilutės atskiri nariai, gali būti paverčiami indeksais arba indeksiniais skaičiais: reikalinga tiktai, kad indeksų arba indeksinių skaičių vaizduojamieji statistikos duomenys būtų lyginami su kitu tos pačios, arba kitos, statistinės eilutės skaičiumi arba kuriuo nors būdu sujungtų skaičių vidurkiu. Skaičius, kuriuo indeksuojamas statistikos duomuo dalomas, yra vadinamas indekso (indeksinio skaičiaus) baze. Sulig tuo, ar indeksuojamos statistinės eilutės skaičiai dalomi vienu ir tuo pačiu skaičiumi, — ar kuriuo nors kriterijum bazinis skaičius keičiamas, — statistikos indeksai, visų pirma, skirstomi į dvi pagrindines grupes: a) indeksai su pastovia (nekintama) baze ir b) indeksai su nepastovia arba keičiama (kintama) baze. Tiek pirmoji, tiek antroji indeksų grupė dar gali būti skirstoma į pogrupes sulig tuo, kuriuo principu pastovioji arba keičiamoji (kintamoji) bazė sudaroma.

Be to, statistikos praktikoj indeksiniais skaičiais būna vaizduojamos paimtos atskirai ne tik statistikos eilutės, bet ir statistinių eilučių vienos ar kitos rūšies junginys. Šia prasme gali būti individualiniai ir kolektyviniai indeksai. Jei kolektyviniai indeksai sudaromi neatsižvelgiant indeksų jungiamųjų st. eilutės skaičių vad. lyginamojo svorio, tai jie vadinami paprastais aritmetiniais indeksais, o jei atsižvelgiama — svertiniais indeksais (pastarieji kai kurių autorių vadinami dar agregatiniais indeksais). Be to, gali būti ir vad. matematiniai kolektyviniai indeksai.

Patys indeksiniai skaičiai gali būti pareiškiami, t. y. lyginami su baziniais skaičiais, sveikais (natūriniais) skaičiais arba trupmenomis; be to, sveikieji skaičiai gali būti pareiškiami procentais, promiliais ar kt. Sulig tuo, kurios rūšies santykiniais

skaičiais indeksai sudaromi, bazinis skaičius prilyginamas vienetui, šimtui ar kt. Skaičiavimo teorijos atžvilgiu nesvarbu, ar bazinis skaičius bus prilygintas 1, 10, 100, 1000, 1000000 ar kt. Tačiau statistikos praktikoj bazinis skaičius dažniausia prilyginamas 100, — kad indeksiniai skaičiai būtų procentiniame santykyje su bazės skaičiumi. Daug rečiau jau sutinkami indeksiniai skaičiai, kuriems bazė būna prilyginta 1. Indeksinių skaičių, kuriems bazė būtų prilyginta 10, 1000 ar kt., statistikos praktikoj nesutinkama.

19. Indeksai su pastovia baze.

Pagrindinė indeksų su pastovia baze sąlyga yra ta, kad visi tos pačios st. eilutės indeksiniai skaičiai būtų betarpiškoj pareinamybėj nuo bazinio skaičiaus (visi būtų dalomi tuo pačiu skaičiumi). Tačiau tuomi dar nepasakoma, kad indeksams baze tinka bet kuris skaičius. Atvirkščiai, kaip visur statistikoj, taip ir čia indeksai turi būti statistiškai prasmingi, t. y. atitikti bendruosius statistikos teorijos reikalavimus ir specifinius, jau tik indeksams statomus, uždavinius.

Statistikos teorija reikalauja, kad kiekviena statistinė eilutė būtų didžiųjų skaičių dėsnio veikiama bent tiek, kiek tas veikimas yra būtinas masės arba masinio reiškinio charakteringoms ypatybėms pasireikšti. Ši pati taisyklė galioja ir statistikos indeksams bei jų bazei. Jei norima bet kurios rūšies statistinės eilutės originalius (absoliutinius) skaičius pakeisti indeksiniais (santykiniais) skaičiais su pastovia baze, tai bazę atstovaujantys skaičiai, aplamai imant, turėtų būti kiek galint „normalesni“.

Turint prieš akis šį bendrąjį statistikos dėsnį ir prisiminus, kad statistinės eilutės indeksų bazei reikalingieji skaičiai gali būti imami ir iš tos pačios skaičių eilutės, — statistikos indeksų bazei prašyte prašosi vidurkis.

Tačiau vidurkių statistikoj yra visokių rūšių; taip pat ir metodų vidurkiams surasti yra ne vienas. Kad indeksiniai skaičiai tiktų statistikos apibrėžiamam uždaviniui spręsti, reikalingas atitinkantis vidurkis.

a) Indeksai su baze iš aritmetinio vidurkio.

Šios rūšies indeksai parodys kiekvieno statistinės eilutės nario santykinį dydį (lyginį) su jos aritmetiniu vidurkiu. Jei x-inės eilutės sekoj atskirus narius pažymėsime $x_1, x_2, x_3, \ldots, x_n$, o aritmetinį vidurkį ($\bar{x}$) prilyginsime 1, tat indeksiniai skaičiai i_i seks šitaip:

$$i_1 = \frac{x_1}{\bar{x}}, \quad i_2 = \frac{x_2}{\bar{x}}, \quad i_3 = \frac{x_3}{\bar{x}}, \ldots\ldots, \quad i_n = \frac{x_n}{\bar{x}},$$

o aplamai,

$$i_i = \frac{x_i}{\bar{x}} \qquad \text{(LIII)}.$$

Jei indeksų būtų ieškoma bazinį skaičių prilyginus 100, tat indeksiniams skaičiams ta pati formula LIII rašytina:

$$i_{i\%} = \frac{x_i}{\bar{x}} \cdot 100, \qquad \text{(LIIIa)}$$

o eilutės sekoj eitų indeksai:

$$i_1 = \frac{x_1}{\bar{x}} 100; \quad i_2 = \frac{x_2}{\bar{x}} 100, \ldots\ldots, \quad i_n = \frac{x_n}{\bar{x}} 100.$$

Statistikos praktikoj aritmetinis vidurkis indeksinių skaičių baze, tačiau, tikslinga yra imti tik tuo atveju, kai indeksuojama skaičių eilutė iš anksto aprėžiama ar kai norima kiekvieno st. eilutės nario lyginamąjį svorį su vidurkiu pabrėžti. Bet jei panorėtume, pav., aritmetiniu vidurkiu bazuoti indeksus prekių kainos kitimams, tai kiekvienam sekančiam kainos kitimui (sekančiam mėnesiui ar kt.) reikėtų iš naujo vidurkis surasti, taigi vis nauja bazė — šiuo atveju keičiamasis vidurkis — paimti. O tai jau prieštarautų bazės pastovumo sąvokai ir neatitiktų indeksams su pastovia baze statomus reikalavimus. Dėl šios priežasties statistikos praktikoj indeksai su baze iš aritmetinio vidurkio paprastai yra vartojami sezoniniams ar, aplamai, periodinams svyravimams atvaizduoti arba išmatuoti. Nes jeigu statistinėj eilutėj yra bet kurios rūšies pariodiškai pasikartojantieji svyravimai, tai tiems svyra-

vimams bendras, o visai eilutei charakteringas, matas ir bus aritmetinis vidurkis. Iš anksto gi neaprėžtoms arba tęstinėms statistikos eilutėms indeksuoti (pav., kainų ar kitų reiškinių einamiems kitimams vaizduoti) yra reikalinga pastovi bazė, kuri jau nepareitų nuo kiekvieno naujai prisidedančio (pav., sekančio mėn.) dydžio.

b) Indeksai su baze iš mediano.

Kiekvieną statistinę eilutę tam tikru atžvilgiu charakterizuoja ir medianas arba v i d u r i n i s dydis, kuris, kaip žinoma, vad., „normalioj" eilutėj (st. eilutėj su normalia dispersija) atitinka ir aritmetinį vidurkį. Taigi medianas gali būti visai tinkama bazė statistikos indeksams tais atvejais, kai statistinė eilutė, iš kurios medianinis dydis imamas indeksų bazei, yra pakankamai ilga ir yra pagrindo teigti, kad ir numatomoj ateityje galimi kitimai (atskirų dydžių skirtumai) neišeis iš normalių nukrypimų (vidurinio kvadratingojo nukrypimo) ribos.

Vis dėlto statistikos praktikoj ligi šiol medianinis dydis indeksų bazei imamas, palyginti, retai. Priežastis ta, kad patsai vidurinis griežtai aprėžtos statistinės eilutės dydis vis dėlto daug pareina ir nuo statistinės eilutės atsitiktinai paimtųjų dydžių (narių) skaičiaus. Ir juo mažesnis medianui parinkti paimtos statistinės eilutės narių skaičius ir didesni jos atskirų dydžių skirtumai, — tuo medianinė bazė silpniau apibūdins (charakterizuos) statistinę eilutę.

Indeksai su baze iš M formula būtų šitokia:

$$i_M = \frac{x_i}{M}\,;\quad i_{\%M} = \frac{x_i \cdot 100}{M} \qquad \text{(LIIIb).}$$

c) Indeksai su baze iš dominanto.

Šios rūšies indeksų formulą tenka rašyti šitaip:

$$i_D = \frac{x_i}{D}\,;\quad i_{\%D} = \frac{x_i \cdot 100}{D}\,.$$

Indeksai su baze iš dominanto, atrodo, jau geriau atitinka statistikos teorijos reikalavimus. Nes dominantas, t. y. daž-

niausiasis dydis, prie visokių aplinkybių bus artimesnis vidurkiui arba „normaliam“ dydžiui. Statistikos praktikoj, tačiau, ir dominantas gana retai imamas baze indeksams skaičiuoti. Dalykas toks, kad dominantas, lygiai kaip ir medianas, vis dėlto tikslingiau atstovauti tegali tik tą statistinę eilutę, iš kurios jis yra paimtas. Praktikoj bet kada gali atsitikti, kad jau poros pirmųjų dydžių statistinėj eilutėj prikergimas dominantinį dydį gali pakeisti. Be to, kiekvienoj atskirai paimtoj statistinėj eilutėj dominantas turi s a v ą j į svorį (vienoj statistinėj eilutėj dominantas gali apimti 50% ir daugiau visų eilutės dydžių, o kitoj — gal neapims nė 10%). Ši dominanto ypatybė daro jį netinkamu dydžiu indeksų bazei ypač tais atvejais, kai atskirų statistinių eilučių individualiniai indeksai tenka jungti ir sudaryti vienos ar kitos rūšies kolektyvinius indeksus. O statistikos praktikoj toksai individualinių indeksų jungimas į kolektyvinius indeksus yra nuolatinis reiškinys.

d) Indeksai su baze iš „norminių“ dydžių.

Paprastiems, iš anksto speciališkai aprėžtų uždavinių neturintiems, indeksams pastovią bazę parinkti yra tikslingiausia iš tokių dydžių, kurie būtų kiek galint „normalesni“. Ši sąlyga, nuosekliai vykdoma, turėtų vesti prie bazinio skaičiaus, kuris atstovautų visus tuos statistinės eilutės dydžius, kurie yra normalių arba nuolatinių veiksnių įtakoje. Konkretiškai kalbant, galima samprotauti, kad, pav., jei yra žinoma, jog konjunktūros ciklų dauguma svyruoja 3 ir 10 metų ribose, tai konjunktūros ar, aplamai, ekonominio gyvenimo reiškinių statistikoj naudojamiems indeksams su pastovia baze reikėtų imti atitinkamų st. eilučių normalius dydžius (arba vidurkius) ne mažiau kaip iš 3 ir ne daugiau kaip iš 10 metų stebėjimo.

Statistikos praktikoj gana dažnai baziniai dydžiai iš tokių metų skaičiaus ir parenkami. Pav., anglų „Statist“ arba vad. Sauerbeck indeksams dar ir pokariniais laikais baziniai skaičiai buvo naudojami iš 1867—1877 metų vidurkio. „Economist“ savo prekių kainų indeksams bazei naudoja 1901—1905 m. vidurkius. Tarptautinis Statistikos Institutas savo indeksams bazę imdavo iš 1924—1925 m. vidurkio. Ligi pa-

starųjų laikų daugelio kraštų ir tautų statistikos praktikoj indeksams būdavo naudojama bazė iš „prieškarinių laikų“ ir tenkinamąsi, dažniausia, vienų 1913 ar 1914 m. vidurkiais ne dėl to, kad tokia, palyginti, siaura bazė būtų laikoma ideališka, o dėl to, kad prieškariniams laikams reikalingieji statistikos duomenys ilgesniam periodui gana sunkiai gaunami. Be to, žymiai pakitusios pokarinių metų sąlygos, iš tikrųjų, ir dabar dar gali būti laikomos nenormalėmis, lyginant su sąlygomis, kurios buvo bet kuriais prieškariniais metais. Tik pačiu pastaruoju laiku vad. „didžiosios krizės“ (1929—1936) metų įtakoje pradėta ypač ūkio reiškinius dažniau matuoti ir lyginti su „pokariniais“ metais.

Jei sutarsime vienu ar kitu kriterijum „normaliai bazei“ paimtus sudarytus x-inės eilutės dydžius žymėti raide m, tai atitinkanti indeksavimo formula atrodys šitaip:

$$i_m = \frac{x_i}{m}; \qquad i_{\%m} = \frac{x_i \cdot 100}{m} \qquad \text{(LIIIc)}.$$

e) Indeksai su baze iš kraštutinių dydžių.

Tam tikrais atvejais gali būti tikslinga indeksų bazei paimti bet kurį kraštutinį dydį (pirmąjį ar paskutinį narį eilutės sekoj, arba didžiausiąjį ar mažiausiąjį dydį). Toki indeksai parodys atskirų dydžių santykį su kraštutiniais arba su maksimaliniais nukrypimais. Šį indeksavimo principą taikant konjunktūros statistikai baze galima pasirinkti dydžius iš „geriausiųjų“ (aukštos konjunktūros) arba iš blogiausiųjų (ekonominės depresijos) metų, kad ryškiau pabrėžti konjunktūros kritimą. Pav. Tautų Sąjungos einamosios statistikos leidiniuose pastaraisiais metais indeksams bazė imama, aplamai, iš 1929 m. (paskutinių prieškrizinių metų) vidurkio.

Apibendrinant šiam reikalui indeksavimo formulą galima rašyti:

$$i_k = \frac{x_i}{x_k}; \qquad i_{\%k} = \frac{100\,x_i}{x_k} \qquad \text{(LIIId)}.$$

f) Indeksai su baze iš specialiai parinktų dydžių.

Kai kada statistikos praktikoj indeksų pastoviai bazei reikalingas dydis parenkamas ir specialiai. Pav., gana dažnai

galima rasti indeksuotų statistinių eilučių, kur pirmasis eilutės dydis paimtas visiems sekančiųjų dydžių indeksams baze. Taip pat neretai galima rasti statistinių eilučių, indeksuotų sulig baze iš kurio nors gana trumpo periodo ar net vienodo kurio momento pastebėtų dydžių. Nulemia čia kurie nors statistinės praktikos ar skaičiavimo technikos sumetimai, bet gali būti ir rimti ekonomiškai teoretiniai arba statistiškai metodologiniai samprotavimai. Pav., 1931 m. Anglijai ir kai kurioms kitoms valstybėms atsisakius aukso standarto ir staiga pinigų vertei nukritus, susidarė visai naujos sąlygos kainų kitimams ir kitiems ūkinio gyvenimo reiškiniams matuoti. Norint šiuos naujus ūkinio gyvenimo veiksnius bei reiškinius statistikos metodais atvaizduoti, juo labiau analizuoti, tenka ir indeksiniams skaičiams bazę pritaikinti tam momentui, kuriuo pasibaigė, arba nuo kurio prasidėjo, tam tikrų, ligi tol nebuvusių, aplinkybių arba veiksnių įtaka. Šias sąlygas atitinkančią indeksavimo formulą galima šitaip rašyti:

$$i_s = \frac{x_i}{x_s}; \qquad i_{s\%} = \frac{100\,x_i}{x_s} \qquad \text{(LIIIe).}$$

Aplamai, indeksai su pastovia baze galima šitaip apibrėžti:

1) n o r m a l i e j i i n d e k s a i, skaičiuojami su baze iš normalių dydžių arba iš normalaus periodo — tai dažniausiai ir plačiausiai vartojami indeksai bet kurios rūšies reiškiniams vaizduoti be griežtai apibrėžtų specifinių uždavinių;

2) s p e c i f i n i a i i n d e k s a i su aprėžtais uždaviniais:

a) indeksai su baze iš tos pačios st. eilutės vidurkio — tinkamiausi sezoniniams ir, aplamai, periodiniams svyravimams matuoti arba dispersijos pobūdžiui nustatyti;

b) indeksai su baze iš kraštutinių dydžių;

c) indeksai su baze iš specialiai parinktųjų dydžių.

20. Indeksai su nepastovia (keičiama) baze.

Indeksai su nepastovia arba keičiamąja baze statistikos praktikoj sutinkami trejopi: a) kintamieji, b) kumuliatyviniai (jungtiniai) ir c) periodiniai indeksai.

a) Kintamieji indeksai.

Kintamieji indeksai, kaip jau patsai pavadinimas rodo, privalo nuolat k i n t a n č i o s bazės. Tokią bazę gali sudaryti bet kuriuo principu parinktas tos pačios st. eilutės skaičius taip, kad kiekvienas indeksuojamas skaičius gautų tuo pačiu kriterijum parinktą, bet tik sau, t. y. nuo kitų skirtingą bazę. Statistikos praktikoj, tačiau, kintamieji indeksai paprastai skaičiuojami visuomet su baze iš ankstyvesniojo dydžio tuo būdu, kad apibendrinta formula šitaip atrodo:

$$i_{k,i} = \frac{x_i}{x_{i-1}} \qquad \text{(LIV).}$$

Eilutės sekoj, kur atskirai paimtųjų skaitomųjų periodų (metų, mėnesių ar kt.) duomenys eitų kaip x_1, x_2, x_3, ... x_n kintamieji indeksai sudaromi šitaip:

$$i_{k,1} = \frac{x_1}{x_0},\ i_{k,2} = \frac{x_2}{x_1},\ i_{k,3} = \frac{x_3}{x_2},\ \ldots,\ i_{k,n} = \frac{x_n}{x_{n-1}}.$$

Tačiau, jeigu eilutės sekoj, kuri prasideda x_1, kintamos bazės sąlygas atitinkančio x_0 nebūtų, tai tektų pirmąjį skaičių palikti neindeksuotą, arba duoti jam sutartą bazinį dydį (1 arba 100), o toliau jau eitų kaip ir anksčiau

$$(i_{k,2} = \frac{x_2}{x_1} \text{ ir t. t.}).$$

Patiektieji pavyzdžiai rodo, kad kintamieji indeksai betarpiškai susiejami tik su gretimaisiais st. eilutės dydžiais, ir tik gretimai esantieji indeksiniai skaičiai parodo jų savitarpio lyginamąjį svorį. Ši kintamųjų indeksų ypatybė daro juos tinkamais tik specifiniams st. analizio uždaviniams, pav., kai norima indeksais pavaizduoti, kiek procentų kurio nors skaitomojo periodo (metų, mėnesio ir kt.) stebimas reiškinys pakitėjo (padidėjo ar sumažėjo) lyginant su ankstyvesniuoju periodu. Tačiau, statistikos praktikoj kintamieji indeksai dažniausia yra tik priemonė aukštesniojo laipsnio indeksams (ypač vad. normalių periodų arba vad. normaliai sezoniniams indeksams) sudaryti.

b) Jungtiniai kumuliatyviniai indeksai.

Kumuliatyviniai indeksai dar daugiau kaip kintamieji indeksai tarnauja tik priemone aukštesniojo laipsnio indeksams sudaryti. Be to, kumuliatyviniai indeksai ir iš viso tėra skaičiuojami ne betarpiškai originaliniams st. eilutės skaičiams pakeisti, o vienu ar kitu metodu sudarytų indeksinių skaičių eilutei „sujungti". Statistikos praktikoj paprastai „kumuliuojami" kintamieji indeksai pagal šią formulą:

$$j_{k(i)} = \frac{x_1}{x_0} \cdot \frac{x_2}{x_1} \cdot \frac{x_3}{x_2} \cdots \frac{x_i}{x_{i-1}} \cdots \frac{x_n}{x_{n-1}}$$

$$j_{k(i)} = i_{k,1} \cdot i_{k,2} \cdots i_{k,i} \cdots i_{k,n} \cdot \qquad \text{(LV).}$$

Kumuliatyviniai indeksai, kaip ir kintamieji indeksai, dažniausia yra vartojami sudarant vad. normaliai sezoninius indeksus.

c) Periodiniai bei sezoniniai indeksai.

Periodiniai bei sezoniniai indeksai, arba indeksai su periodine (sezonine) baze, privalo parodyti bet kurio reiškinio santykį su atitinkamo periodo (sezono) charakteringomis ypatybėmis. Pav., jei kuris nors socialinio ar ekonominio gyvenimo reiškinys yra susietas su mažiau ar daugiau reguliariai pasikartojančių veiksnių įtaka, tai iš indeksinių skaičių su bet kurios rūšies pastovia baze dar nebus galima spręsti, ar tam tikro meto kitimai pareina nuo tų periodiškai pasikartojančių veiksnių, ar nuo kitų kurių priežasčių. Imkime, pav., kiaušinių kainas. Rudenį ir žiemą tos kainos paprastai būna aukštesnės negu pavasarį ar vasarą. Čia grynai sezoninis reiškinys, pareinąs nuo vištų didesnio dėslumo pavasario mėnesiais. Bet kiaušinių, kaip ir kiekvienos kitos prekės, kainos pareina ne tik nuo pasiūlų, bet ir nuo paklausų. O paklausai yra veikiami ir bendrųjų vad. ūkio konjunktūros sąlygų. Taigi tam tikru metu pastebėtas kiaušinių kainos kilimas ar kritimas gali pareiti ir ne nuo sezoninių veiksnių (vištų dėslumo), ar net ne nuo konjunktūrinių, o nuo struktūrinių (paklausų sumažėjimo arba padidėjimo, vištų laikymo ir priežiūros, išlaidų padidėjimo arba sumažėjimo ir t. t.). Taip pat ir sezoninių ar, apla-

mai, periodinių veiksnių įtaka dėl įvairių priežasčių gali pasikeisti. Visus šios rūšies kitimus tegali atvaizduoti tik indeksai su periodine baze.

Periodinę bazę indeksams parenkant galioja bendrasis indeksų bazės parinkimo dėsnis: ir periodinė bazė turi būti t i p i š k a arba „normaliai“ periodinė, t. y. kiek galint mažiau pareiti nuo atsitiktinių veiksnių įtakos. Sulig indeksams statomais uždaviniais, periodinius indeksus galima suskirstyti dar į šias pogrupes: 1) paprastųjų periodų indeksai, 2) aprėžtųjų periodų indeksai ir 3) normaliųjų periodų indeksai.

1. Paprastųjų periodų indeksai.

Paprastųjų periodų indeksai sudaromi tais atvejais, kai norima kurio nors skaitomojo periodo duomenis pavaizduoti indeksiniais skaičiais lyginant su tuo pačiu skaitomuoju periodu iš bet kurio kito laikotarpio (pav., einamųjų metų atskirų mėnesių, importo, eksporto, kainų, gimimų ar kt. duomenys lyginami su atitinkamų mėnesių tais pačiais duomenimis iš praeitųjų, užpraeitųjų ar kt. metų).

Jei mes sutarsime indeksuojamo skaitomojo periodo bet kurios rūšies duomenis žymėti x_i, o praeitojo ar kt. lyginimui imamo periodo $x_{\bar{i}}$, tai paprastųjų periodų indeksiniai skaičiai bus gaunami pagal formulą:

$$i_{\bar{i}} = \frac{x_i}{x_{\bar{i}}} \qquad \text{(LVI).}$$

Eilutės sekoj atskiriems periodams indeksai išeis:

$$i_{\bar{i}} = \frac{x_1}{x_{\bar{i},1}},\quad i_{\bar{i},2} = \frac{x^2}{x_{\bar{i},2}},\quad x_{\bar{i},3} = \frac{x_3}{x_{\bar{i},3}},\quad \ldots,\quad i_{\bar{i},n} = \frac{x_n}{x_{\bar{i},n}}.$$

Žodžiu, kiekvieno skaitomojo periodo duomuo indeksuojamas baze iš to paties skaitomojo periodo, paimto, tačiau, iš kito, specialiai lyginimo reikalui numatyto, periodo. Šios rūšies periodiniai indeksai gana dažnai sutinkami einamosios statistikos šaltiniuose.

2. Aprėžtųjų periodų indeksai.

Aprėžtųjų periodų indeksai skaičiuojami su baze iš plačiau aprėžto periodo, negu einamiems indeksams skaitomieji periodai. Nulemti čia gali skaičiavimo technika arba ir teoretiškai metodologiniai sumetimai. Pav., duomenys indeksuojamam laikotarpiui gali būti smulkesniais skaitomaisiais periodais suskirstyti kaip duomenys lyginamajam (bazės) periodui (savaitiniai šalia mėnesinių, mėnesiniai šalia kvartalinių ir t. t.). Bet taip pat gali nulemti ir metodologiniai samprotavimai. Pav., skaitomųjų periodų indeksams bazė gali būti imama iš specifiškai aprėžto laikotarpio, k. a.: mėnesiniai indeksai skaičiuojami atskirai su baze iš pavasario, vasaros, rudens ir žiemos mėnesių vidurkio, arba iš aukščiausiųjų ar žemiausiųjų kainų (didžiausių ir mažiausių pasiūlų) laikotarpio ir kt.

Aprėžtųjų periodų indeksai skaičiuojami pagal formulą:

$$i_{a,i} = \frac{x_i}{\ddot{x}_{\bar{i}}} \qquad \text{(LVII)}.$$

Šioj formuloj indeksuojamas skaičius pažymėtas x_i, o bazės laikotarpiui aprėžtųjų periodų atitinkami skaitomųjų periodų vidurkiai $\ddot{x}$. Eilutės sekoj, kur indeksuojama 12 mėnesių eilutę atstovautų $x_1, x_2, x_3, \ldots, x_n$ ir, aplamai, x_i, o bazės laikotarpio (kitų metų) atitinkamų mėnesių duomenys atžymėti $x'_1, x'_2, x'_3, x'_4, \ldots, x'_m$, ir, aplamai, x'_j, aprėžtųjų, tarsime, trimėnesinių periodų indeksai būtų skaičiuojami šiuo būdu:

$$i_{a,1} = \frac{x_1}{\ddot{x}_{\bar{1}}} = \frac{x_1}{\frac{1}{3}(x'_1 + x'_2 + x'_3)},$$

$$i_{a,2} = \frac{x_2}{\ddot{x}_{\bar{1}}},$$

$$i_{a,3} = \frac{x_3}{\ddot{x}_{\bar{1}}},$$

$$i_{a,4} = \frac{x_4}{\ddot{x}_{\bar{1}|\bar{1}}} = \frac{x_4}{\frac{1}{3}(x'_4 + x'_5 + x'_6)},$$

$$i_{a,5} = \frac{x_5}{\ddot{x}_{II}},$$

.....

.....

.....

$$i_{a,7} = \frac{x_7}{\ddot{x}_{III}} = \frac{x_7}{\frac{1}{3}(x'_7 + x'_8 + x'_9)},$$

.....

.....

.....

$$i_{a,10} = \frac{x_{10}}{\ddot{x}_{IV}} = \frac{x_{10}}{\frac{1}{3}(x'_{10} + x'_{11} + x'_{12})},$$

.....

.....

.....

$$i_{a,12} = \frac{x_{12}}{\ddot{x}_{IV}}.$$

Nesunku pastebėti, kad aprėžtųjų periodų indeksais vaizduojamoji st. eilutė yra, tarytum, suskaldyta į keletą savo tarpe nesurištų grupių. Be to, šiuo metodu parenkamoji bazė atskiriems kitimų ciklo periodams, arba sezonams, lengvai pasiduoda indeksų sudarytojo savivaliai. Nes griežtai vienodi kitimų ciklo periodai jų ilgumo ir kitimų laipsnio atžvilgiais labai retai tegali pasitaikyti. Esant ne visai vienodo ilgumo periodams, arba pasireiškus atskiruose perioduose skirtingiems kitimų laipsniams, griežtai vienalyčių bazinių skaičių sudarymas yra sunkiai įmanomas ir visuomet mažiau ar daugiau pareis indeksuotojo nuožiūros.

Statistikos praktikoj šios rūšies indeksai mažai ir vartojami. Ir Lietuvoj Centralinio Statistikos Biuro veikimo pradžioje metų kainų indeksai šiuo būdu tik kurį laiką būdavo sudaromi.

3. Normalūs periodiniai bei sezoniniai indeksai.

Normalūs periodiniai (sezoniniai) indeksai statistikos praktikoj vartojami tais atvejais, kai norima kuris nors reiškinys atvaizduoti nepriklausomai nuo sistematiškai (periodiškai) pasikartojančių veiksnių įtakos, arba kai norima tų veiksnių (sezoninių svyravimų) įtaką eliminuoti. Statistiškai metodologiniu atžvilgiu pirmoj vietoj čia stovi ne patys indeksiniai skaičiai, o metodai, kurie nulemia vad. normaliųjų periodų arba normaliųjų sezonų bazę. Tų metodų moderniškoji konjunktūros statistika gana daug turi ir vis dar naujų išranda. Jiems visiems čia net ir labai trumpai apibrėžti perdaug vietos užimtų, dėl to pasitenkinsime tik tais, kurie šių laikų statistikos praktikoj dažniau sutinkami ir kurie bet kuriuo metu gali būti reikalingi.

Kaip jau minėta, normalių periodinių arba normalių sezoninių indeksų statistinę vertę nulemia tikslingai bei prasmingai sudaryti baziniai skaičiai, šiuo atveju — normali periodinė bazė.

Ši normalioji periodinė bazė yra reikalinga duomenų iš ilgesnio laikotarpio. Kiek ilgo laikotarpio skaitomųjų periodų duomenys normaliai bazei yra reikalingi, — pareina nuo periodiniais indeksais vaizduotinos statistinės eilutės ypatybių. Aplamai, ir čia, kaip ir visur panašiais atvejais statistikoj, galioja dėsnis: juo didesni yra statistinės eilutės dydžių skirtumai — tuo ilgesni periodai reikalingi normaliosios bazės skaičiams gauti. Ir atvirkščiai: esant nedideliems statistinės eilutės dydžių skirtumams, galima tenkintis trumpesnio periodo skaičiais. Vis dėlto prie jokių aplinkybių nerekomenduojama normaliai periodinei bazei imti vidurkinius dydžius iš mažiau kaip trejeto vad. ciklų.

Statistikos praktikoj normaliai periodinei bazei reikalingieji skaičiai dažniausia imami iš 5—10 atitinkamų ciklų arba laikotarpių (pav., 5—10 metų duomenys normaliems sezoniniams indeksams sudaryti). Jei dėl vienų ar kitų priežasčių pakankamai ilgam periodui st. duomenų nėra, o statistinės eilutės dydžių skirtumai nėra maži, kai kada rekomenduojama vienu ar kitu metodu tokią statistinę eilutę pirmiau išlyginti, t. y. atsitiktinius nukrypimus eliminuoti ir tik paskui normaliai periodinius arba normaliai sezoninius indeksus sudaryti.

d) Periodinių vidurkių metodas normaliai sezoniniams indeksams skaičiuoti.

Paprasčiausias metodas normaliai sezoninių arba periodinių indeksų bazei sudaryti — tai paimti keletą ciklų apimančio laikotarpio skaitomųjų periodų vidurkius. Jei mes sutarsime periodiniams (sezoniniams) svyravimams numatytus ciklus (metus) eilutės sekoj žymėti $x_I, x_{II}, \ldots, x_m$ ir, aplamai, x_j, o to paties ciklo atskirų metų atskirai paimtus skaitomuosius periodus (mėnesius) jų eilutės sekoj žymėsime $x_1, x_2, x_3, \ldots, x_n$, aplamai, x_i, tai normaliai periodinę bazę sudarys visų ciklų (metų) Σy_j atitinkamų periodų (mėnesių) x_i duomenų vidurkis $\bar{x}_j$, kurį galima atžymėti:

$$\bar{x}_j = \frac{\Sigma x_j}{m} = \frac{x_I + x_{II} + \ldots + x_m}{m} \qquad \text{(LVIII).}$$

Šiuo būdu kiekvieno skaitomojo periodo (mėnesio) x_i dydis dalomas iš ciklų (metų) eilės išvestųjų vidurkių $\bar{x}_j$, kurių turi būti tiek, kiek yra x_i-nės eilutės skaitomųjų periodų (mėnesių), t. y. $\Sigma \bar{x}_j = 12$. Jei sutarsime ciklų (metų) skaičius m apibrėžti 5 ir turėsime penkių metų mėnesinius duomenis, tai gausime šitokią schemą: (žiūr. psl. 375).

Gautieji atskirų mėnesių (periodiniai) vidurkiai $\bar{x}_j$ duos normalų, kiekvieną skaitomąjį periodą (šiuo atveju mėnesį) atitinkantį, dydį, kuriame normaliai sezoninių veiksnių įtaka bus pasireiškusi, o atsitiktinių veiksnių įtaka eliminuota. Periodiniai (atskirų mėnesių) vidurkiai $\bar{x}_j$ gali būti panaudoti normaliai periodiniams (sezoniniams) indeksams, kuriais sezoninių veiksnių įtaka gali būti eliminuota. Šitai atliekama tuo būdu, kad gautieji atskiri periodiniai vidurkiai pirmiau indeksuojami pagal formulą:

$$i_s = \frac{\bar{x}_j}{\bar{x}_{ij}} = \frac{\frac{1}{m}\Sigma x_j}{\frac{1}{n}\Sigma \bar{x}_i} = \frac{\Sigma x_j}{\frac{1}{n}\Sigma\Sigma x_{ij}} \qquad \text{(LVIX).}$$

Formula rodo, kad šiuo metodu normaliai sezoninis indeksas gaunamas arba betarpiškai iš originalių skaičių atitinkamų periodinių dydžių bendrųjų skaičių, juos indeksuojant baziniu

Lentelė 49.

PERIODINIŲ VIDURKIŲ METODO SCHEMA.

I.	x_1,	x_2,	x_3,	x_4,	x_5,	x_6,	x_7,	x_8,	x_9,	x_{10},	x_{11},	x_{12}	$\Sigma x_{i,I}$
II.	x_1,	x_2,	x_3,	x_4,	x_5,	x_6,	x_7,	x_8,	x_9,	x_{10},	x_{11},	x_{12}	$\Sigma x_{i,II}$
III.	x_1,	x_2,	x_3,	x_4,	x_5,	x_6,	x_7,	x_8,	x_9,	x_{10},	x_{11},	x_{12}	$\Sigma x_{i,III}$
IV.	x_1,	x_2,	x_3,	x_4,	x_5,	x_6,	x_7,	x_8,	x_9,	x_{10},	x_{11},	x_{12}	$\Sigma x_{i,IV}$
V.	x_1,	x_2,	x_3,	x_4,	x_5,	x_6,	x_7,	x_8,	x_9,	x_{10},	x_{11},	x_{12}	$\Sigma x_{i,V}$
	Σx_j	Σx_I,	Σx_{II},	Σx_{III},	Σx_{IV},	Σx_V,	Σx_{VI},	Σx_{VII},	Σx_{VIII},	Σx_{IX},	Σx_X,	Σx_{XI},	Σx_{XII}	$\Sigma\Sigma x_{ij}$	
	$\frac{1}{m}\Sigma\bar{x}_j$:	$\bar{x}_I$,	$\bar{x}_{II}$,	$\bar{x}_{III}$,	$\bar{x}_{IV}$,	$\bar{x}_V$,	$\bar{x}_{VI}$,	$\bar{x}_{VII}$,	$\bar{x}_{VIII}$,	$\bar{x}_{IX}$,	$\bar{x}_X$,	$\bar{x}_{XI}$,	$\bar{x}_{XII}$	$\bar{\Sigma x_j} = \frac{\Sigma\Sigma x_{ij}}{m}$	

skaičium iš jų vidurkio, arba iš keleto metų periodinių (mėnesinių) vidurkių $\bar{x}j$. Šiuo būdu gautieji normaliai sezoniniai indeksai jau gali būti naudojami absoliutinių skaičių sezoniniams svyravimams eliminuoti pagal formulą:

$$x'_1 = \frac{x_1}{i_{s,1}}, \quad x'_2 = \frac{x_2}{i_{s,2}}, \quad x'_3 = \frac{x_3}{i_{s,3}}, \ldots, x'_n = \frac{x_n}{i_{s,n}},$$

o, aplamai:

$$x'_i = \frac{x_i}{i_{s,i}} \qquad \text{(LX)}.$$

Tačiau, šiuo metodu išskaičiuotas normaliai sezoninis indeksas nevisuomet yra tinkamas sezoninių veiksnių įtakai eliminuoti. Būtent, tais atvejais, kai normaliai periodinei bazei sudaryti paimtame laikotarpyje yra buvusi kurios nors pašalinės kilmės (nepaprastų aplinkybių arba struktūrinio pobūdžio) veiksnių įtaka, tai ir normaliai sezoniniame indekse, sudarytame periodinių vidurkių metodu, šitoji įtaka bus palikusi. O tada, eliminuojant periodinių vidurkių metodo indeksu sezoninius svyravimus, būtų eliminuota ne tik sezoninių veiksnių įtaka, bet ir struktūriniai, arba net ir konjunktūriniai, kitimai.

e) Periodiškai kolektyvinių indeksų metodas.

Šis metodas reikalauja viso numatytojo laikotarpio atskirų ciklų skaitomųjų periodų (mėnesių) duomenis iš karto indeksuoti su baze iš tų pačių ciklų vidurkio. Paskui gi, susumavus atskirų ciklų tų pačių skaitomųjų periodų indeksinius skaičius, jau išskaičiuoti normaliai sezoninį indeksą su baze iš jų pačių vidurkio.

Jei normaliai sezoniniam indeksui sudaryti numatytojo laikotarpio atskirų sezonų ciklus (metus), kaip ir anksčiau, žymėsime $x_I, x_{II}, \ldots, x_m$ ir, aplamai, x_j, o atskirai paimtų ciklų skaitomųjų periodų (mėnesių) duomenis eilutės sekoj $x_1, x_2, x_3, \ldots, x_n$, aplamai, x_i, tai normaliai sezoniniam indeksui sudaryti bus reikalinga:

1. Kiekvieno sezonų ciklo (metų) x-inei eilutei sudaryti indeksus su baze iš jos pačios vidurkio pagal formulą LIII, kur:

$$i_{\bar{x},i} = \frac{x_i}{\bar{x}} = \frac{x_i}{\frac{1}{n}\Sigma x_i}.$$

Lentelė 50.

NORMALIAI SEZONINIS INDEKSAS PERIODINIŲ VIDURKIŲ METODU (ARITMETINIS).

Nr.	Metai mj / Mėnesiai (n,)	Kiaušinių mėnesinės kainos (x_j) litais 1932—1936 m. (x_j)					1932—36 m mėnesiais sk. periodų x_j sumos (Σx_j)	Mėnesiais skait. periodų (m) vidurkiai: $\bar{x}_j = \frac{1}{m}\Sigma x_j$	Normaliai sezoninis indeksas $i_s = \frac{\bar{x}_j}{\bar{x}_{jj}} = \frac{\frac{1}{m}\Sigma x_j}{\frac{1}{n}\Sigma \bar{x}_j}$
		1932	1933	1934	1935	1936			
	0	1	2	3	4	5	6	7	8
1	Sausis . . .	1,40	1,15	1,40	1,05	0,85	5,85	1,17	1,37
2	Vasaris . . .	1,20	1,20	0,90	1,05	0,90	5,25	1,05	1,23
3	Kovas . . .	1,25	0,80	0,75	0,60	0,80	4,20	0,84	0,98
4	Balandis . . .	0,75	0,65	0,60	0,60	0,70	3,30	0,66	0,77
5	Gegužė . . .	0,70	0,65	0,55	0,55	0,60	3,05	0,61	0,71
6	Birželis . . .	0,70	0,65	0,65	0,55	0,60	3,15	0,63	0,73
7	Liepa	0,80	0,65	0,65	0,55	0,60	3,25	0,65	0,76
8	Rugpiūtis . .	0,80	0,70	0,65	0,55	0,60	3,30	0,66	0,77
9	Rugsėjis . . .	0,95	0,80	0,65	0,60	0,70	3,70	0,74	0,86
10	Spalis. . . .	1,05	1,05	0,75	0,70	0,80	4,35	0,87	1,02
11	Lapkritis. . .	1,55	1,25	0,90	0,90	1,20	5,80	1,16	1,35
12	Gruodis . . .	1,50	1,40	1,05	1,00	1,23	6,18	1,24	1,45
	Iš viso . . .	12,65	10,95	9,50	8,70	9,58	$\Sigma\Sigma x_{ji} = 51,38$	$\Sigma \bar{x}_j = 10,28$	$\Sigma i_s = 12,00$

Lentelė 51.

NORMALIAI SEZONINIS INDEKSAS PERIODIŠKAI KOLEKTYVINIŲ INDEKSŲ (PERIODOGRAFINIU) METODU.

Nr.	Metai (mj) / Mėnesiai (nj)	Kiaušinių kainos indeksai su baze iš vidurkio: $i_x = \frac{x_i}{\bar{x}}$					Mėnesinių indeksų suma (Σi_j)	Normaliai sezoninis indeksas $i_s = \frac{\frac{1}{m}\Sigma i_j}{\frac{1}{n}\Sigma i_i}$
		1932 (i_1)	1933 (i_2)	1934 (i_3)	1935 (i_4)	1936 (i_5)		
	0	1	2	3	4	5	6	7
1	Sausis . . .	1,33	1,26	1,77	1,45	1,07	6,88	1,38
2	Vasaris . .	1,14	1,32	1,13	1,45	1,13	6,17	1,23
3	Kovas . . .	1,19	0,88	0,95	0,83	1,00	4,85	0,97
4	Balandis. .	0,71	0,71	0,76	0,83	0,88	3,89	0,78
5	Gegužė . .	0,66	0,71	0,70	0,76	0,75	3,58	0,72
6	Birželis . .	0,66	0,71	0,82	0,76	0,75	3,70	0,74
7	Liepa . . .	0,76	0,71	0,82	0,76	0,75	3,80	0,76
8	Rugpiūtis .	0,76	0,77	0,82	0,76	0,75	3,86	0,77
9	Rugsėjis . .	0,90	0,88	0,82	0,83	0,88	4,31	0,86
10	Spalis . . .	1,00	1,15	0,95	0,96	1,00	5,06	1,01
11	Lapkritis .	1,47	1,37	1,13	1,23	1,50	6,70	1,34
12	Gruodis . .	1,42	1,53	1,33	1,38	1,54	7,20	1,44
		12,00	12,00	12,00	12,00	12,00	60,00	12,00

2. x_j ciklams gautuosius atskirų periodų (mėnesių) indeksus $i_{\bar{x},i}$ sumuoti ir padalinti periodų (m) skaičiumi:

$$i_j = \frac{i_{\bar{x},I} + i_{\bar{x},II} + \ldots + i_{\bar{x},m}}{m} = \frac{1}{m}\Sigma i_{\bar{x},j} .$$

3. Gautuosius atitinkamų mėnesių indeksų vidurkius eilutės sekoj sudėti ir gauti jų vidurkį:

$$\bar{i} = \frac{\Sigma i_j}{n} = \frac{1}{n}(i_1 + i_2 + i_3 + \ldots + i_n).$$

4. Šis indeksų vidurkis $\bar{i}$ sudarys bazę normaliai sezoniniam indeksui pagal formulą:

$$i_s = \frac{i_j}{\bar{i}} = \frac{\frac{1}{m}\Sigma i_{\bar{x},j}}{\frac{1}{n}\Sigma i_j} \qquad \text{(LXI).}$$

Gautieji normaliai sezoninio indekso i_s skaičiai naudojami, kaip ir paprastu aritmetiniu gautieji, sezoniniams svyravimams eliminuoti tuo pačiu būdu, kaip parodyta psl. 376 (pagal formulą LX).

i) Persons'o metodas.

Trečiasis metodas normaliai sezoniniams indeksams yra siejamas su Persons'o vardu. Statistikos praktikoj naudojama keletas šio metodo variantų. Tačiau visų jų pagrindinė esmė yra ta, kad Persons'o metodas ligi tam tikro laipsnio įgalina iš normaliai periodinės bazės paimtojo laikotarpio nustatyti ir eliminuoti ypatingųjų sąlygų, arba vad. struktūrinės kilmės, sudarytąją specifinę kryptį ir tuo pačiu paryškinti konjunktūros kitimus.

Persons'o metodu sudaromieji normaliai sezoniniai indeksai reikalingi šių operacijų:

1. Normaliajai bazei paimtojo laikotarpio mėnesiniai duomenys indeksuojami su kintamąja baze pagal formulą LIV, kur

$$i_{k,i} = \frac{x_i}{x_{i-1}}$$

2. Viso laikotarpio atskirų metų atitinkamų mėnesių kintamieji indeksai sudedami ir dalomi metų skaičiumi pagal formulą LVIII, kur:

$$\bar{i}_{k,i} = \frac{1}{m}\Sigma i_{k,j} = \frac{i_{k,I} + i_{k,II} + \ldots + i_{k,m}}{m}$$

3. Gautieji kintamųjų indeksų periodiniai (mėnesiniai) vidurkiai jungiami vad. multiplikatyviniu metodu į kumuliatyvinius indeksus pagal LV formulą:

$$j_i = \bar{i}_{k,1} \cdot \bar{i}_{k,2} \cdot \ldots \cdot \bar{i}_{k,n}$$

Eilutės sekoj

$$j_1 = \bar{i}_1,\ j_{k,2} = \bar{i}_1 \cdot \bar{i}_2,\ j_3 = \bar{i}_1 \cdot \bar{i}_2 \cdot \bar{i}_3 \ldots,\ j_n = \bar{i}_1 \cdot \bar{i}_2 \cdot \bar{i}_3, \ldots,\ \bar{i}_n.$$

4. Jeigu, sudauginus, eilutės sekoj paskutinis indeksas j_n skiriasi nuo 1, tatai parodo, kad normaliai periodiniam indeksui paimtame laikotarpyje būta struktūrinės kilmės veiksnių, kurie, jei stipriau pasireiškia, reikėtų eliminuoti. Tai daroma tuo būdu, kad paskutiniojo (gruodžio) mėnesio indeksas (j_{12}) atimamas nuo vieneto ir gautasis skirtumas, padalinus mėnesių skaičiumi, atimamas arba pridedamas atitinkamu eilutės sekos skaičiumi sudauginus — sulig tuo ar $1 - j_{12}$ skirtumas gaunamas teigiamas arba neigiamas. Tuo būdu koreguotieji kumuliatyviniai indeksai gaunami pagal formulą:

$$j'_i = j_i + \delta x = j_i + \frac{1}{12}(1 - j_{12})x,$$

Kur x yra natūrinis aritmetinės eilutės skaičius: $x_1, x_2, \ldots, x_{12}$. Tuo būdu gauname:

$$j'_1 = j_1 + \frac{1}{12}(1 - j_{12})1,$$

$$j'_2 = j_2 + \frac{1}{12}(1 - j_{12})2,$$

$$j'_3 = j_3 + \frac{1}{12}(1 - j_{12})3,$$

$$\ldots\ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots$$

$$j'_{12} = j_{12} + \frac{1}{12}(1 - j_{12})\,12 = 1.$$

5. Šiuo būdu koreguotieji kumuliatyviniai indeksai j'_i jau naudojami periodiškai sezoniniam indeksui kaip ir ankstybesnieji (pagal formulas LIX arba LXI sudarytieji)

$$i_{s,p} = \frac{j'}{\bar{j}} = \frac{j'_i}{\frac{1}{n}\Sigma j_i}. \qquad \text{(LXII).}$$

Normaliai sezoniniai indeksai i_s, gautieji Persons'o metodu, naudojami sezoninių veiksnių įtakai eliminuoti taip pat pagal formulą LX:

$$x'_i = \frac{x_i}{i_{s,i}}.$$

Lentelė 52.

NORMALIAI SEZONINIS INDEKSAS PERSONS'O METODU.

Nr.	1932 — 36 metų kiaušinių kainos kintamieji indeksai $i_k = \frac{x_i}{x_{i-1}}$						Kintamųjų indeksų sumos (Σi_k)	Kintamųjų indeksų vidurkiai $\left(\frac{1}{m}\Sigma i_k = i_x\right)$	Jungtiniai (kumuliatyviniai) indeksai ($j = i_1 \ldots .. i_i \ldots i_{12}$)	Koreguoti jungtiniai indeksai $\left(j = j_i + \frac{1-j_{12}}{12} x\right)$	Normaliai sezoninis indeksas $\left(i_s = \frac{j_i}{\frac{1}{12}\Sigma j_i}\right)$
	Mėnesiai	1932	1933	1934	1935	1936					
	0	1	2	3	4	5	6	7	8	9	10
1	Sausis . . .	1,00	0,77	1,00	1,00	0,85	4,62	0,92	0,92	0,91	1,34
2	Vasaris . .	0,86	1,04	0,64	1,00	1,06	4,60	0,92	0,85	0,82	1,21
3	Kovas . . .	1,04	0,67	0,83	0,57	0,89	4,00	0,80	0,74	0.70	1,03
4	Balandis . .	0,60	0,81	0,80	1,00	0,88	4,09	0,82	0,61	0,55	0,81
5	Gegužė . .	0,93	1,00	0,92	0,92	0,86	4,63	0,93	0,57	0,50	0,73
6	Birželis. . .	1,00	1,00	1,18	1,00	1,00	5,18	1,04	0,59	0,51	0,75
7	Liepa . . .	1,14	1,00	1,00	1,00	1,00	5,14	1,03	0,61	0,51	0,75
8	Rugpiūtis. .	1,00	1,08	1,00	1,00	1,00	5,08	1,02	0,62	0,51	0,75
9	Rugsėjis . .	1,19	1,14	1,00	1,09	1,17	5,59	1,12	0,69	0,56	0 82
10	Spalis . . .	1,11	1,31	1,15	1,17	1,14	5,88	1,18	0,81	0,67	0,99
11	Lapkritis . .	1,48	1,19	1,20	1,29	1,50	6,66	1,33	1,08	0,92	1,35
12	Gruodis . .	0,97	1,12	1,17	1,11	1,03	5,40	1,08	1,17	1,00	1,47
		12,32	12,13	11,89	12,15	12,38	60,87	12,19	9,26	8,16	12,00

Lentelė 53.

NORMALIAI SEZONINIŲ INDEKSŲ PALYGINIMAS IR PERIODINIŲ VIDURKIŲ INDEKSO SEZONINIŲ SVYRAVIMŲ ELIMINAVIMAS.

Nr.	Mėnesiai	Periodiniai vidurkiai ($\bar{x}_j$)	Norm. periodiniai indeksai			1936 mt. kiaušinių kainos (x_i)	Paprastas sezoninis indeksas su baze iš period. vidurkių $\left(i_s = \frac{x_i}{\bar{x}_j}\right)$	Kiaušinių kaina eliminavus sezoninius svyravimus	
			($i_{\bar{x},j}$)	($i_{\bar{i}}$)	(i_p)			$\left(\ddot{x}_i = \frac{x_i}{i_{\bar{x},j}}\right)$	$\left(\ddot{x}_i = \frac{x_i}{i_p}\right)$
	0	1	2	3	4	5	6	7	8
1	Sausis . . .	1,17	1,37	1,38	1,34	0,85	0,73	0,62	0,63
2	Vasaris . . .	1,05	1,23	1,23	1,21	0,90	0,86	0,73	0,74
3	Kovas . . .	0,84	0,98	0,97	1,03	0,80	0,95	0,82	0,78
4	Balandis . . .	0,66	0,77	0,78	0,81	0,70	1,06	0,91	0,86
5	Gegužė . . .	0,61	0,71	0,72	0,73	0,60	0,98	0,85	0,82
6	Birželis . . .	0,63	0,73	0,74	0,75	0,60	0,95	0,82	0,80
7	Liepa	0,65	0,76	0,76	0,75	0,60	0,92	0,79	0,80
8	Rugpiūtis .	0,66	0,77	0,77	0,75	0,60	0,91	0,78	0,80
9	Rugsėjis . . .	0,74	0,86	0,86	0,82	0,70	0,95	0,81	0,85
10	Spalis. . . .	0,87	1,02	1,01	0,99	0,80	0,92	0,78	0,81
11	Lapkritis. . .	1,16	1,35	1,34	1,35	1,20	1,03	0,89	0,89
12	Gruodis . . .	1,24	1,45	1,44	1,47	1,23	0,99	0,85	0,84
		10,28	12,00	12,00	12,00	9,58	11,25	9,65	9,62

Psl. 377—378 ir 381—382 yra patiektos lentelės, kur normaliai sezoniniai indeksai yra išskaičiuoti visais trimis metodais kiaušinių kainoms. Lentelėje 53 visi trys normaliai sezoniniai indeksai yra sugretinti (žiūr. 2, 3 ir 4 skilt.), be to, pirmojoj skiltyj yra pakartoti iš lent. 52, 7 sk. periodiniai vidurkiai, o iš tos pačios lentelės 5 sk. yra patiektos 1936 mt. kiaušinių kainos. 6 sk. patiekti paprastieji sezoniniai indeksai su normaliai periodine baze, o sk. 7 — šio indekso pagalba koreguotos, t. y. nuo sezoninių veiksnių įtakos eliminuotos, kiaušinių kainos. Be to, sk. 8 yra patiektos tos pačios kainos, eliminavus iš jų sezoninius svyravimus Persons'o indeksu.

Palyginus visais trim metodais sudarytus normaliai sezoninius indeksus (lent. 53, sk. 2—4) galima įsitikinti, kad esminių skirtumų tiems patiems skaičiams nė vienas metodas nesudaro. Net ir reliatyviniai skirtumai nedideli (daugumą sudaro 1—2 punktai) ir didžiausias skirtumas, kurį sudaro periodinių kolektyvinių indeksų ir Persons'o metodai sausio mėn. tesudaro tik 4 p. (138 : 134).

Atrodytų, kad painesnieji ir daugiau darbo reikalingieji metodai visai nereikalingi. Tačiau „Socialekonominės Statistikos" psl. 521—532 demonstruotuose pavyzdžiuose galima rasti skirtumų, kurie siekia keliolika punktų. O tai jau gali turėti didelę įtaką sezoninius svyravimus eliminuojant. Priežastis ta, kad panaudotuose duomenyse struktūriškai konjunktūrinio pobūdžio veiksnių įtaka buvo nevienoda. Tai matyti iš Persons'o metodo veikimo: kai „Socialekonominėj Statistikoj" panaudotieji duomenys sviesto eksportui kilimo tendenciją gruodžio mėn. jungtiniame indekse pareiškė nukrypusią nuo normos (I) net 81 p. ($j_{12} = 1{,}81$), tai kiaušinių kainose ši įtaka aprėžiama tik 17 p. ($j_{12} = 1{,}17$).

Šie pavyzdžiai pirštu prikišamai parodo Persons'o metodo tikrąją reikšmę ir statistinę prasmę. Aišku, kad vieno ar kito metodo pasirinkimas pareina nuo pageidaujamo griežtumo laipsnio. Aplamai, tačiau, kai normaliai sezoniniam indeksui paimtame laikotarpyje struktūrinio ar konjunktūrinio pobūdžio veiksnių įtaka (vienašališka kitimų tendencija) vieną kurią laikotarpio dalį yra atitolinusi nuo normos bent 20—25%, griežtesni metodai jau daug gali nulemti. O statistine būtinybe Persons'o ar kt. griežtesnis metodas virsta tais atvejais, kai ši

tendencija pasireiškia stiprumu, siekiančiu 50% į vieną ar kitą pusę. Pažinti tendencijos stiprumo laipsnį apytikriai galima iš atskirų ciklų (metų) vidurkių. Pav., jeigu paimti 1932 ir 1933 m. vidutines kainas iš vienos pusės ir 1935—1936 m. kainas iš antrosios, tai rasime, kad skirtumas sudaro apie 13%.

Normaliai sezoniniams indeksams paimtojo periodo tendenciją iš dalies galima eliminuoti dar ir kraštutinius, t. y. mažiausius ir didžiausius kiekvienų metų dydžius išbraukiant. Tuo būdu iš lentelės 52 patiektųjų duomenų, pirmiau negu sumuoti tektų išbraukti: pirmojoj skiltyj (1932 m.) gegužės ir lapkričio mėn. kainas, antrojoj (1933 m.) balandžio ir gruodžio mėn. kainas ir t. t. Kraštutinių dydžių eliminavimas gali būti prasmingas net ir Persons'o metode, — ypač tais atvejais, kai normaliai sezoniniam indeksui naudojama ne daugelio ciklų (metų) duomenys, ar kai juose būna nepaprastai dideli atsitiktiniai nukrypimai.

Kai kurie kiti normaliai sezoninių indeksų sudarymo variantai yra paminėti „Socialekonominėj Statistikoj" (skyriuje apie konjunktūros statistiką).

21. Kolektyviniai indeksai.

a) Kolektyvinių indeksų esmė ir jų sudarymo sąlygos.

Atskirai paimtoms statistinėms eilutėms, k. a. bet kurios prekės kainai, bet kurios gėrybės ar produkto gamybai, eksportui, importui ar kt. išskaičiuotieji indeksai yra individualiniai indeksai, nes jie atstovauja tik vieną ar vienalytę statistinę eilutę.

Kolektyviniai indeksai, kaip rodo ir patsai vardas, atstovauja statistikos eilučių vienu ar kitu būdu sujungtą kolektyvą. Pav., atskirų javų (rugių, kviečių, miežių ir kt.) kainoms išskaičiuotieji individualiniai indeksai gali būti vienu ar kitu metodu jungiami į kolektyvinį javų indeksą. Pieno, kiaušinių, sviesto ir kt. atskirai paimtų maisto produktų kainoms indeksai gali būti jungiami į kolektyvinį maisto produktų indeksą ir t. t. Konjunktūros statistikoj jungiami į kolektyvinius indeksus ir visai negimininingų statistinių eilučių

indeksai, k. a., kainų indeksai su derlių ar kitų gėrybių gamybos indeksais, eksporto ir importo indeksai su kainų, akcijų kurso, pinigų apyvartos ir kt. indeksais, ir t. t.

Pagrindinė sąlyga kolektyviniams indeksams prasmingais būti — tai tos pačios rūšies bazė. Kolektyviniams indeksams gali būti imami ir jungiami tik tokie ir tik tie individualiniai indeksai, kuriems skaičiuoti buvo panaudota vienoda, t. y. tuo pačiu principu parinkta bazė. Negalima sudaryti kolektyvinių indeksų statistinėms eilutėms, iš kurių vienoms individualiniai indeksai būtų buvę sudaryti su pastovia baze, o kitoms — su kintamąja; arba vienoms su pastovia baze iš aritmetinio vidurkio, o kitoms — iš mediano ar kurio kito dydžio.

Antroji, tam tikromis aplinkybėmis nemažiau svarbi, kolektyvinių indeksų statistinio prasmingumo sąlyga — tai pačių statistinių eilučių, arba jas atstovaujančių individualinių indeksų, charakterio (kitimų) kilmės giminingumas. Jei kolektyviniam indeksui būtų jungiamos statistinės eilutės, iš kurių vienos būtų veikiamos periodiškai pasikartojančių veiksnių ir turėtų, pav., charakteringus sezoninius svyravimus, o kitos sezoninės kilmės svyravimų neturėtų, arba tie svyravimai visai kitus periodus turėtų, tai iš tokių statistinių eilučių sudarytas kolektyvinis indeksas būtų grynai aritmetinio skaičiavimo padarinys — be statistinės prasmės. Norint kolektyvinius indeksus statistiškai prasmingus turėti, reikalinga viena iš dviejų: a) periodinių (sezoninių) veiksnių įtaką iš originalinių skaičių, arba juos atstovaujančių individualinių skaičių, eliminuoti arba b) kolektyviniams indeksams jungti tik vienodo charakterio periodinių svyravimų individualinius indeksus.

b) Kolektyvinių indeksų sudarymo metodai.

Kolektyviniai indeksai gali būti: 1) paprastieji aritmetiniai indeksai, 2) svertiniai indeksai arba ir 3) vad. matematiniai indeksai.

a) Paprastieji aritmetiniai indeksai individualinių indeksų kolektyvams (prekių grupių indeksams, vad. generaliniam indeksui ir t. t.) gaunami lygiai taip, kaip paprас-

tieji aritmetiniai vidurkiai: individualiniai indeksai sumuojami ir paskui dalomi sumos skaičiumi pagal VIII formulą. Tačiau, kad nepainioti vidurkių su indeksais, geriau yra iš karto ir simbolius pakeisti. O kadangi mes jau įpratome individualinius indeksus žymėti raide i, tai bus visai logiška, jei kolektyvinį indeksą žymėsime I, ir tuo būdu gausime atitinkamai pakeistą formulą:

$$I = \frac{1}{n} \Sigma i = \frac{i_1 + i_2 + i_3 + \ldots + i_n}{n} \qquad \text{(LXIII)}.$$

Šiuo metodu, pav., mūsų Centr. Stat. Biuras patiekia tiek atskiroms prekių grupėms indeksus, tiek bendrąjį visoms prekių kainoms (generalinį) indeksą „Statistikos Biuletenyje".

b) Svertiniai indeksai individualinių indeksų kolektyvams, kaip jau pats vardas rodo, yra reikalingi svorio koeficientų. Jie ir skaičiuojami pasikliaujant svertinių vidurkių (VIIIa) formula. Tačiau, kai svertiniams vidurkiams svorio koeficientai paprastai būna betarpiškai susiję su vidurkiais vaizduojamų originalinių dydžių dispersija ir visuomet yra lygūs tų dydžių atstovaujamiems skaičiams vienoj statistinėj eilutėj, tai indeksams svorio koeficientai yra reikalingi specifinio parinkimo.

c) Svorio koeficientų indeksams parinkimo sąlygos.

Statistikos teorija reikalauja, kad statistikos metodais vaizduojama masė, arba masiniai reiškiniai, atitiktų charakteringas ypatybes. Indeksiniais skaičiais, paprastai, vaizduojami masiniai reiškiniai (prekių kainos, materialinių gėrybių gamyba, prekių eksportas, darbininkų darbo algos ir t. t.). Bet tie reiškiniai gyvenimo tikrenybėje nėra lygiareikšmiai. Pav., masinio vartojimo prekių, k. a.: duonos, sviesto, bulvių, druskos ar kt. kainos ir jų kitimai tautos ūkiui turi žymiai didesnės reikšmės, kaip specifinio ar prabangos prekių (brangiųjų gėrimų, ekzotinių kraštų vaisių ar kt. maisto produktų ar p.) kainos ir jų kitimai. Taip pat vienos ūkinės reikšmės indeksais vaizduojami javų ar gyvulių ūkio gamybos kitimai, kitos reikšmės tautos ūkiui turi durpių arba plytų gamyba, dar kitos —

arbatos, degtinės arba sėmenų aliejaus gamyba. Net ir tos pačios ūkinės kategorijos reiškiniai gali nevienodą ūkinę vertę turėti. Pav., druska ir duona yra masinio vartojimo prekės, bet atskiro asmens arba šeimos biudžetui 10% kainos pakilimas duonai ir 10% kainos pakilimas druskai yra toli gražu ne tos pačios reikšmės (pav., Lietuvos Centr. Stat. Biuro skelbiamiems gyvenimo minimumo indeksams 1 asmeniui išlaidos druskai sudaro tik apie 0,1%, o išlaidos duonai beveik 7%). Taip pat ir tautos ūkiui vienos reikšmės bus sumažėjimas trečdaliu ar 50% linų, bekono arba sviesto eksporto, o kitos — tokiu pat koeficientu sumažėjimas ožkų, kiškių arba šieno eksporto. Tuo būdu ir kolektyviniai indeksai tik tuomet statistikos teorijos statomus reikalavimus gali patenkinti, kai jie sudaryti atsižvelgiant tautos ūkyje esamo lyginamojo svorio tų reiškinių, kuriuos individualiniai indeksai atstovauja.

Statistiškai metodologiniu atžvilgiu individualinių indeksų kolektyviniam indeksui atitinkantis „svėrimas" ir svorio koeficientas juo yra reikalingesnis, juo kolektyviniam indeksui naudojamų individualinių indeksų skaičius mažesnis ir juo didesni yra jų skirtumai ir, žinoma, atvirkščiai: juo daugiau individualinių indeksų kolektyviniam indeksui naudojama ir juo mažesni jų skirtumai, — juo mažesnę statistinę reikšmę turės ir bet koki svorio koeficientai.

Sunkiausia svorio koeficientus nustatyti prekių kainų kolektyvams (prekių grupėms) ir bendrajam kainolygiui. Ideališkas bet kurios prekių grupės ar vad. bendrojo šalies kainolygio indeksas turėtų būti sudarytas atsižvelgiant kiekvienos atskirai paimtos prekės lyginamojo svorio t a u t o s ū k y j e. Bet kokiais ir kuriais objektyviniais pažymiais tasai ū k i n i s s v o r i s statistikai nustatytinas?

Pirmiausia prašyte prašosi svorio koeficientams g a m y b o s k i e k i s. Bet gamybos statistinis stebėjimas yra gana komplikuotas, o daugeliui prekių net neįmanomas. Be to, yra prekių, kurios šalies viduje visai negaminamos, bet jų apyvarta bei vartojimas gali prašokti ir daugelį šalies viduje gaminamų gėrybių.

Ši aplinkybė iškelia mintį, kad svorio koeficientams tikslingiau gali būti imami v a r t o j i m o k i e k i a i. Bet vartojimo statistinis stebėjimas yra nemažiau komplikuotas,

kaip gamybos, — ypač jeigu paimti gėrybes, kurios naudojamos asmeniniam ir ūkiniam vartojimui ir priklauso vad. elastingųjų paklausų dėsnio. Pav., rugiai yra vartojami sėklai, duonai ir pramonės reikalams. Tačiau pastoviausi paklausai tėra tik sėklai. Rugių paklausai duonai taip pat gana pastovūs, nors ir ne tiek, kaip sėklai. O rugių paklausai pramonės reikalams (spiritui varyti) pareina nuo derliaus ir kainos ne tik rugių, bet ir jų pakaitalų (bulvių arba kitų duoninių javų). Tuo būdu ir vartojimo kiekiai, kaip dažnai kintantieji dydžiai, mažai tinka pastoviems svorio koeficientams; juo labiau, kad patys kitimai nesiduoda statistikos metodais griežtai apibrėžiami.

Atrodytų, kad svorio koeficientams geriausia tiktų prekių a p y v a r t a. Tačiau apyvartos statistinis stebėjimas nemažiau komplikuotas, kaip gamybos arba vartojimo. Svarbiausia, yra nemaža prekių, kurių apyvarta rinkoje negali būti tos prekės lyginamojo svorio tautos ūkyje rodikliu. Pav., Lietuvoje rugių auginama apie keturius kartus daugiau, kaip kviečių. Bet jeigu kuriais metais derlius gaunamas tik vidutinis, tai beveik visas tų pačių ūkininkų suvartojamas duonai bei sėklai. O tokiose aplinkybėse ir rugių kainos padidėjimas arba sumažėjimas tautos ūkyje ir pačių ūkininkų bus jaučiama daug mažiau, kaip kviečių kainos tas pats pakitimas: nes kviečių derliaus dauguma visuomet parduodama. Pav., jeigu patys ūkininkai sėklai ir maistui rugių yra reikalingi per metus apie 450 000 tonų, o kviečių tik 50 000 t, tai iš 500 000 t rugių derliaus rinkai teks tik 50 000 t, o iš 150 000 t kviečių derliaus — net 100 000 t. Savaime aišku, kad prie tokių ar panašių aplinkybių 2 litais rugių ir kviečių kainos pakitimas ir rinkoje perkančių rugių vartotojų ir jų gamintojų bus mažiau jaučiamas, kaip kviečių vartotojų ir gamintojų, — nežiūrint rugių didesnės gamybos ir didesnio bendrojo vartojimo. Bet jeigu paimsime druską, silkes ar kurią kitą šalies viduje negaminamą — arba kad ir gaminamą, bet pačių gamintojų betarpiškai nevartojamą prekę (k. a. audinius, batus ir p.), — tai kaina, kuri nusistato rinkos apyvartoj sulig paklausais ar pasiūlais, vienodai bus jaučiama ir tos prekės tiekėjų, ir jos vartotojų. Iš to matyti, kad s t a t i s t i š k a i p r a s m i n g i s v o r i o k o e f i c i e n t a i t u r i b ū t i n u s t a t o m i s u l i g k i l m ė s i r s p e-

cifinių vartojimo sąlygų kiekvienai prekei atskirai. Kainos prekių, kurios vartotoją pasiekia tik per tarpininkus iš rinkos, bendrąjį kainolygį ir ūkio konjunktūrą turi veikti daugiau, kaip kainos prekių, kurios vartojamos pačių gamintojų betarpiškai. Dėl to ir tikslinga, kad bendrąjį kainolygį vaizduojantiems kolektyviniams indeksams svorio koeficientai būtų atitinkamai pritaikinti.

Jau šie pavyzdžiai ir statistiškai metodologiniai samprotavimai rodo, kad kolektyviniams indeksams statistiškai prasmingų svorio koeficientų nustatymas yra nelengvas ir, gal būt, ideališkai net neišsprendžiamas uždavinys. Paprastos statistinės eilutės aritmetiniams vidurkiams svorio koeficientai visuomet bus ideališki, jei tik jie atitiks kiekvieno atskirai paimto dydžio lyginamąjį svorį visoj statistikos eilutėj, kurios narių skaičius yra griežtai aprėžtas. Visai kitos aplinkybės yra indeksinių skaičių eilutėse, kurios tuo ir pasižymi, kad statistinės prasmės jos įgauna tik kitimų variantuose tęstinėj (vis augančioj) st. eilutėj. O jeigu taip, tai gali atsitikti, kad vienam periodui nustatyti visais atžvilgiais ideališki svorio koeficientai kitu metu, ar kitose aplinkybėse, gali ne tik metodologiško tikslumo bei statistinio prasmingumo nustoti, bet, gal būt, ir tikrenybę iškreipti.

d) Individualinių indeksų svėrimo sistemos.

Svertinių indeksų istorija taip pat rodo, kad indeksų svėrimo metodai yra labai įvairūs ir, svarbiausia, ligi šiol statistikos praktikoj nėra vyraujančio (dominuojančio) svėrimo metodo, — nežiūrint į tai, kad patys indeksai, kaip statistinio vaizdavimo priemonė, jau turi apie 80 metų amžiaus.

Pirmykštė, o dar ir šiandien plačiausia, indeksinio vaizdavimo sritis yra kainų statistikoj. Pritaikė jį Laspeyres 1864 m. vokiečių „Jahrbuecher fuer Nationaloekonomie und Statistik", pasiremdamas Soetbeer'o kainų statistinio stebėjimo duomenimis ligi 1856 m. Laspeyres paėmė 48 prekių poros dešimčių (ligi 1851 m.) vidutines kainas bazei, kuriai prilygino tų pačių prekių kainas visiems sekantiems metams (ligi 1862 m.). Laspeyres darbą vėliau (1874 m.) tęsė

Paasche, tik prekių skaičių sumažino (ligi 22) ir kiek pakeitė metodą.

Dalykas toks, kad tiek Laspeyres, tiek Paasche atskirų prekių kainos kitimus vaizdavo ne individualiniais indeksais, kaip šiandien paprastai daroma, o iš karto kolektyviniu kainų rodikliu. Ši aplinkybė juos vertė iš karto ieškoti svorio koeficientų, t. y. atskirai paimtoms prekėms kurių nors kiekių. Šie kiekiai kiekvienai prekei buvo nustatyti sulig spėjamo vartojimo arba apyvartos dydžiu. Skirtumas tas, kad Laspeyres atskirų prekių kainas vaizdavo kiekiams, paimtiems sulig menamu vidutiniu vartojimu iš indeksų bazei paimtojo periodo, o Paasche tuos kiekius ėmė iš stebimojo periodo, kuriam indeksiniai skaičiai nustatomi. Jei prekių kainas pažymėsime raide p, o jų kiekius raide q, ir jeigu indeksų bazės skaitomąjį periodą pažymėsime raide o, o stebimąjį — raide i, tai indeksams skaičiuoti formulos atrodys šitaip:

$$I_L = \frac{\Sigma p_i\ q_0}{\Sigma p_0\ q_0} \qquad \text{(LXIV).}$$

$$I_P = \frac{\Sigma p_i\ q_i}{\Sigma p_0\ q_i} \qquad \text{(LXV).}$$

Jei sutarsime, kad turime sudaryti kolektyvinį indeksą prekėms a, b, c, ..., z, kurių kainos kitimai p laiko bėgyje atžymimi kaip p_0, p_1 p_2, ..., p_n, o, aplamai, p_i, be to, jei vartojimas (apyvarta, gamyba) q tiems patiems skaitomiems periodams atstovaujamas simbolių q_0, q_1, q_2, ..., q_n, ir, aplamai, q_i, tai kolektyvinis indeksas bet kuriam laikotarpiui pagal Laspeyres formulą LXIV bus nustatomas šitaip:

$$I_{i(L)} = \frac{p_i\ q_{0(a)} + p_i\ q_{0(b)} + \ldots + p_i\ q_{0(z)}}{p_0\ q_{0(a)} + p_0\ q_{0(b)} + \ldots + p_0\ q_{0(z)}} = \frac{\Sigma p_i\ q_0}{\Sigma p_0\ q_0} \qquad \text{(LXIV).}$$

Tuo tarpu toms pačioms sąlygoms Paasche formula LXV bus reikalinga šitokio išvystymo:

$$I_{i(P)} = \frac{p_i\ q_{i(a)} + p_i\ q_{i(b)} + \ldots + p_i\ q_{i(z)}}{p_0\ q_{i(a)} + p_0\ q_{i(b)} + \ldots + p_0\ q_{i(z)}} = \frac{\Sigma p_i\ q_i}{\Sigma p_0\ q_i} \qquad \text{(LXV).}$$

Formulos rodo, kad eilutės sekoj gaunamieji kolektyviniai indeksai I_i tik tuomet sutaptų, jei panaudotųjų prekių

apyvarta visą laiką būtų vienoda, t. y. visais atvejais $q_0 = q_i$. Priešingu atveju kolektyvinių indeksų skirtumai eilutės sekoj neišvengiami.

Statistiškai metodologiniu atžvilgiu Laspeyres panaudotas svėrimo metodas yra tuo patogesnis, kad vieną kartą nustatyti svorio koeficientai gali būti naudojami visiems bet kada indeksuojamos eilutės skaičiams. Tuo pačiu palengvinamas ne tik statistinis vaizdavimas, bet ir st. stebėjimas, kuris indeksuojamam periodui aprėžiamas vien kainos kitimų stebėjimu, nekreipiant dėmesio į pasiūlus ar jų apyvartą. Praktikoje šiuo metodu ir dabar dažniausia svertiniai indeksai skaičiuojami.

Galimas taip pat svertinis indeksas su pastoviais svorio koeficientais baziniam periodui, ir su kintamais stebimiems arba vaizduojamiems periodams sulig formula:

$$I = \frac{\Sigma p_i \, q_i}{\Sigma p_0 \, q_0} \qquad \text{(LXVI).}$$

Kai kurie kiti šio metodo variantai yra aprašyti mano „Socialekonominėj statistikoj“ psl. 521—535.

Kolektyvinius indeksus Laspeyres, Paasche ir kt. senesniais laikais skaičiuodavo betarpiškai iš originalinių skaičių, atskirų st. eilučių nepaversdami individualiniais indeksais. Tačiau kolektyvinių indeksų skaičiavimas betarpiškai iš originalinių statistikos eilučių kolektyvo gali turėti stambių trūkumų tais atvejais, kai griežtai tikslių svorio koeficientų gauti negalima (o tai atsitinka dažniausia) ir, be to, patys st. eilučių indeksuotieji dydžiai labai nevienodi. Pav., norint sudaryti kolektyvinį indeksą gyvulių kainoms, tenka turėti reikalo su galvijų ir arklių kainomis, žymimomis tik litais. Čia visai menkas arklio ar karvės kainos pakitimas, arba nedidelė svorio koeficiento paklaida kolektyviniame indekse, jei jis skaičiuojamas betarpiškai iš absoliutinių kainos kitimų, gali visai nustelbti arba eliminuoti kainos kitimus smulkiųjų gyvulių, kurių lyginamasis svoris tautos ūkyje gali būti nemažas. O to jau negali atsitikti, jei kolektyviniai indeksai skaičiuojami iš individualinių indeksų (žiūr. lentelės 54abc, pp. 392, 393 ir 394).

Lentelė 54a.

KAINOS IR JŲ SVORIO KOEFICIENTAI.

Eilės Nr.	Prekės pav.	Kainos kitimai litais (p_i)			Svorio koeficientai (q_i)		
		(p_0)	(p_1)	(p_2)	(q_0)	(q_1)	(q_2)
n	0	1	2	3	4	5	6
1	Arklys	250	200	287	10	8	15
2	Karvė	150	120	160	10	12	5
3	Kiaulė	30	30	40	26	25	15
4	Avis	15	20	16	2	1	5
5	Višta	3	4	2	1	2	5
6	Antis	3	5	3	1	2	5
	Iš viso	$\Sigma p_0 = 451$	$\Sigma p_1 = 379$	$\Sigma p_2 = 508$	$\Sigma q_0 = 50$	$\Sigma q_1 = 50$	$\Sigma q_2 = 50$
	$I_{0/0}$	$\frac{\Sigma p_0}{\Sigma p_0} = 100$	$\frac{\Sigma p_1}{\Sigma p_0} = 84{,}0$	$\frac{\Sigma p_2}{\Sigma p_0} = 112{,}6$			

e) Kolektyvinių indeksų svorio koeficientų apibrėžimas.

Indeksams svorio koeficientų apibrėžimas yra reikalingas visais tais atvejais, kai svorio koeficientai palieka statistinio metodo būtinybė vidurkiniams dydžiams. Kai kada indeksuojamieji dydžiai savuosius svorio koeficientus turi „iš prigimimo“, ir tenka juos tik panaudoti. Pav., indeksiniais skaičiais vaizduojant užsienių prekybos atskirų prekių grupių kitėjimus, svorio koeficientais virsta atskirų prekių vertės (arba kiekio) lyginamasis svoris. Gyvenimo brangumo arba ekzistencminimumo indeksams natūriniai koeficientai yra atskirų gyvenimo reikmenų (maisto ir kt.) kiekiai, kurie būna baziniam ir skaitomajam periodui pastovūs, o kinta tik gyvenimo reikmenų (maisto produktų, buto nuomos ir kt.) kainos. Taigi gyvenimo brangumo indeksams pilnai galioja ir pirmykštė Laspeyres kolektyvinių svertinių indeksų formula.

Statistikos praktikoj, tačiau, nevisuomet taip lengva tobulus indeksų svorio koeficientus apibrėžti.

Tipiškai rinktiniai indeksai pradėti vartoti A. Sauerbeck'o 1886 m. anglų statistikų draugijos žurnale

Lentelė 54b.

SVERTINIŲ INDEKSŲ SKAIČIAVIMO SCHEMA BETARPIŠKAI IŠ KAINŲ KITIMO PAGAL LENTELĖS 54a DUOMENIS.

Eilės Nr.	Pavadinimai	Prekių kainos ir svorio koeficientų sandaugos pagal tris variantus: $p_i q_o$, $p_o q_i$ ir $p_i q_i$.						
		Sandaugos pagal formulą: $p_i q_i$			Sandaugos pagal formulą: $p_o q_i$		Sandaugos pagal formulą: $p_i q_i$	
		(p_0q_0)	(p_1q_0)	(p_2q_0)	(p_0q_1)	(p_0q_2)	(p_1q_1)	(p_2q_2)
	0	1	2	3	4	5	6	7
1	Arklys	2500	2000	2870	2000	3750	1600	4305
2	Karvė	1500	1200	1600	1800	750	1440	800
3	Kiaulė	780	780	1040	750	450	750	600
4	Avis	30	40	32	15	75	20	80
5	Višta	3	4	2	6	15	8	10
6	Antis	3	5	3	6	15	10	15
	$\Sigma p_i q_i$	$4816 = \Sigma p_0 q_0$	$\Sigma p_1 q_0 = 4029$	$\Sigma p_2 q_0 = 5547$	$\Sigma p_0 q_1 = 4577$	$\Sigma p_0 q_2 = 5055$	$\Sigma p_1 q_1 = 3828$	$\Sigma p_2 q_2 = 5810$
	Indeksai: $I_{0/0}$	$\frac{\Sigma p_0 q_0}{\Sigma p_0 q_0} 100 = 100$	$\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} = \frac{4029.100}{4816} (=) 83,7$	$\frac{\Sigma p_2 q_0}{\Sigma p_0 q_0} = \frac{5547.100}{4816} (=) 115,2$	$\frac{\Sigma p_0 q_1}{\Sigma p_0 q_1} 100 = 100$	$\frac{\Sigma p_0 q_2}{\Sigma p_0 q_2} 100 = 100$	$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} = \frac{3828.100}{4577} (=) 83,6$	$\frac{\Sigma p_2 q_2}{\Sigma p_0 q_2} = \frac{5810.100}{5055} (=) 114,9$

Lentelė 54c.

ARITMETINIAI IR SVERTINIAI KOLEKTYVINIAI INDEKSAI IŠ INDIVIDUALINIŲ INDEKSŲ.

Eilės Nr.	Pavadinimai	Individualūs aritmetiniai indeksai $i = \frac{p_i}{p_0}$			Individualinių indeksų sandauga su pastoviais koeficientais ($i_i q_0$)		Individualinių indeksų sandaugos su kintamais koeficientais ($i_i q_i$)	
		$i_{0/0} = \frac{100p_0}{p_0}$	$i_1 = \frac{100p_1}{p_0}$	$i_2 = \frac{100p_2}{p_0}$	($i_1 q_0$)	($i_2 q_0$)	($i_1 q_1$)	($i_2 q_2$)
n	0	1	2	3	4	5	6	7
1	Arklys	100	80,0	114,8	800,0	1148,0	640,0	1722,0
2	Karvė	100	80,0	106,7	800,0	1067,0	960,0	533,5
3	Kiaulė	100	100,0	133,3	2600,0	3465,8	2500,0	1999,5
4	Avis	100	133,3	106,7	266,6	213,4	133,3	533,5
5	Višta	100	133,3	66,7	133,3	66,7	266,6	333,5
6	Antis	100	166,7	100,0	166,7	100,0	333,4	500,0
	$\Sigma i_i p_i$	100	693,3	628,2	4766,6	6060,9	4833,3	5622,0
	$1_{0/0}$	$\frac{\Sigma i_0}{n} = 100$	$\frac{\Sigma i_1}{n} = \frac{693,3}{6}$ (=)115,6	$\frac{\Sigma i_2}{n} = \frac{628,2}{6}$ (=) 104,7	$\frac{\Sigma i_1 q_0}{\Sigma i_0 q_0} = \frac{4766,6}{50}$ (=) 95,3	$= \frac{\Sigma i_2 q_0}{\Sigma i_0 q_0} = \frac{6060,9}{50}$ (=) 121,2	$= \frac{\Sigma i_1 q_1}{\Sigma i_0 q_i} = \frac{4833,3}{50}$ (=) 96,7	$= \frac{\Sigma i_2 q_2}{\Sigma i_0 q_2} = \frac{5622,0}{50}$ (=) 112,4

„Journal of Royal Statistical Society“ ir ligi šiol naudojami anglų ekonominio savaitraščio vardu „S t a t i s t“. A. Sauerbeck'as išrinko 45 prekes, kurios jam atrodė svarbesnės ir kainų kitimo atžvilgiu Anglijos ūkiui tipiškesnės. Visas tas prekes jis suskirstė į šešias nelygias grupes jų ūkinio gimininguмo atžvilgiu tuo būdu, kad būtų galima dar sudaryti dvi pagrindines prekių klases sulig tuo, ar tos prekės vartojamos žmonių maistui, ar kitiems reikalams. Pačius svorio koeficientus kiekvienai prekių grupei Sauerbeck'as nustatė sulig vidurkių, gautų iš tos grupės prekių bendrosios kainos. Praktikoj Sauerbeck'o pirmykštė svorio koeficientų sudarymo schema šitaip atrodo:

SAUERBECK'O TIPIŠKAI RINKTINIŲ INDEKSŲ SUDARYMO SCHEMA 1885 METAMS.

Prekių grupės	Prekių skaičius	Prekių bendroji kaina	Kainų indeksas (vidurkiai)
1. Augalų maistas	8	546	68
2. Guvulių produktai	7	618	88
3. Cukrus (2 rūš.), kava, arbata	4	250	63
I. Maisto produktai	19	1414	74
4. Mineralai	7	463	66
5. Manufaktūra	8	521	65
6. Kitos įvairios prekės	11	836	76
II. Pramoninės gamybos prek.	26	1820	70
Iš viso	45	3234	72

Savo indeksinių skaičių bazei Sauerbeck'as paėmė 10-ies metų (1867—1877) kainų vidurkius, kuriuos prilygino 100. Vėliau S t a t i s t'o baziniams skaičiams metų periodas buvo pakeistas.

„The Economist'o“ tipiškai rinktiniai indeksai nuo „Statist'o“ rinktinių indeksų skiriasi tuo, kad „Economist'as“ savo parinktoms prekių grupėms ėmė kainas ne specialiai (savivalingai) parinktiems dydžiams (mato vienetų tam tikrą skaičių),

o didmenų prekyboj vartojamiems mato vienetams. Be to, prekių grupėms tenkančias kainos bendrąsias sumas baziniams metams padarė savo rūšies svorio koeficientais bendrąjam rodikliui, kuris vėl baziniams metams prilyginamas 100, kad generalinį (kolektyvinį) indeksą gauti.

Iš visų prekių savo indeksams „Economist" paėmė 44 ir jas suskirstė į 5 grupes šiuo būdu:

I. Javai ir mėsa (įeina 10 prekių),
II. Kiti maisto produktai ir skanėstai (6 prekės),
III. Manufaktūra (10 prekių),
IV. Mineralai (8 prekės) ir
V. Kitos įvairios prekės (10 prekių).

Baziniai skaičiai iš karto sudaryti iš 1845—1850 metų kainų vidurkio, vėliau paimta 1901—1905 metų vidurkis, po karo pasekta kitų kraštų indeksus ir bazei paimta 1913 m. kainos.

Lietuvos Banko prekių indeksams svorio koeficientai yra sudaryti atsižvelgiant gamybos ir rinkos apyvartos toms prekėms, kurios šalies viduje gaminamos ir vien rinkos apyvartos importuojamoms prekėms. Importuojamoms prekėms svorio koeficientu paimta jų lyginamoji importo vertė, kur milijonas litų prilyginta koeficiento vienetui. Šalies viduje gaminamoms prekėms svorio koeficientams paimtas gamybos ir rinkos apyvartos kiekių geometrinis vidurkis sulig formulos $k = \sqrt{ab}$, kur a atstovauja gamybą (bendrąją gamybos vertę), o b — rinkos apyvartą (viskas taip pat skaitoma milijonais litų).

Praktikoj ir vietos gamybos prekėms svorio koeficientai sudaryti skirtingai, — sulig tuo, ar tų prekių kainos Lietuvoje betarpiškai veikiamos pasaulinės rinkos kainų, ar pasaulinės rinkos įtaka dirbtinių ekonominės politikos priemonių eliminuojama. Pasaulinės rinkos įtaka laikoma įrodyta, jei tos pačios prekės (ekonominio vartojimo prasme), kurios šalies viduje gaminamos, yra bent kiek žymesniais kiekiais eksportuojamos arba importuojamos. Kai kurioms prekėms, kurių viena dalis nuolatos suvartojama betarpiškai pačių gamintojų ūkiuose (javai ir kt. žemės ūkio produktai) ir

vidaus rinkon, palyginti, tenka nedideli kiekiai (linai ir kt.) eksporto bei importo lyginamoji vertė paimta jų rinkos apyvartos rodikliu. Prekėms gi, kurios pasaulinės rinkos kainų neveikiamos, k. a. monopolinės ir akcizuojamos prekės, imta vidaus rinkos kainos ir kiekiai atsižvelgiant normalinio (vidutinio) vartojimo. Be to, tais atvejais, kai kainos statistinis stebėjimas apima tik vieną kurios nors prekės rūšį (pav., smulkusis cukrus, pušinės lentos ir t. t.), gamybos ir rinkos apyvartos dydžiai atitinkamai koreguoti (padidinti), kad jų svorio koeficientai galėtų atstovauti bendruosius tos prekės gamybos bei rinkos apyvartos kiekius.

Lietuvos Banko svertinių indeksų schema šitaip atrodo:

Lentelė 55.

PREKIŲ GRUPIŲ KAINŲ INDEKSAI (1926—29 m. = 100).

	Svorio koefic.	1930	1931	1932	1933	1934	1935	1936	1937	1938 I	1938 II	1938 III	1938 IV	1938 V
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
I. Javai, linai ir kt. . . (9 prekės)	8	62	47	40	40	40	39	37	46	43	43	43	42	42
II. Gyvuliai ir paukščiai (6 prekės)	5	95	75	47	42	40	30	40	49	48	49	50	51	52
III. Odos, kailiai ir viln. (5 prekės)	1	87	67	35	34	38	38	50	60	56	54	51	50	49
IV. Gyv. maist. prod. . (10 prekės)	6	93	75	55	47	42	35	38	44	49	47	46	46	46
V. Kiti maist. prod. . . (5 prekės)	3	90	87	89	88	87	85	74	74	74	74	74	74	74
VI. Gėrimai ir taboka . (3 prekės)	3	84	85	86	70	67	60	55	55	54	54	54	54	54
VII. Kuras . . . . . (4 prekės)	2	89	80	75	66	61	55	54	57	60	60	60	61	61
VIII. Manufaktūra . . . . (5 prekės)	4	88	74	64	52	50	45	44	45	45	45	45	46	45
IX. Statybos medžiaga . (7 prekės)	3	95	87	72	57	53	49	47	60	63	62	62	62	63
Bendrasis indeksas .	35	84	72	59	52	50	45	45	51	52	51	51	51	51

f) Matematiniai indeksai.

Statistikos praktikoj iš matematikos teikiamųjų metodų kolektyviniams indeksams skaičiuoti naudojami yra beveik išimtinai vad. g e o m e t r i n i a i i n d e k s a i. Geometriniams indeksams skaičiuoti formula galioja mums jau pažįstamoji XV formula geometriniams vidurkiams skaičiuoti. Tik kai geometriniai indeksai skaičiuojami ne sulig absoliutiniais individualinių dydžių (kitimų) skaičiais, o sulig individualiniais indeksais (o taip daroma dažniausia), tai formulos simboliai atitinkamai pakeičiami ir geometrinis indeksas skaičiuojamas:

$$I_g = \sqrt[n]{i_1 \cdot i_2 \ldots i_n}, \text{ arba: } \log I = \frac{\lg i_1 + \lg i_2 + \ldots + \lg i_n}{n} \quad \text{(LXVII).}$$

Jei mes prisiminsime visa tai, kas savo laiku buvo išdėstyta bei išrodyta apie geometrinius vidurkius, kaip priemonę statistinėms eilutėms charakterizuoti, tai iš anksto teks paabejoti geometrinių indeksų tinkamumu kolektyviniams indeksams skaičiuoti. Juk geometrinis vidurkis visais atvejais išeina mažesnis už aritmetinį vidurkį, ir skirtumas yra juo didesnis, juo statistinės eilutės dydžiai nevienodesni. Skyriuje apie statistikos vidurkius mes turėjome progos įsitikinti, kad geometrinis vidurkis visa savo prigimtimi gali būti tinkamesnis charakterizuoti tik tam tikro tipo asimetrinei statistikos eilutei, — kur mažesnieji dydžiai gausingiau atstovaujami. Bet jeigu statistinėj eilutėj gausingiau atstovaujami stambesnieji dydžiai, tai geometrinis vidurkis bus blogesnis jos reprezentantas, kaip paprastas aritmetinis vidurkis. Absoliutiniai gi indeksuojamų individualinių dydžių skaičiai ir patys individualiniai indeksai vienoda tikimybe gali, pav., vieną mėnesį turėti dominantą (gausingiausią dydį) kairėj nuo mediano pusėj, o kitą — dešinėj. O šiose sąlygose viso kolektyvo vaizdavimas vienu atveju gali būti tikslingesnis bei statistiškai prasmingesnis geometriniu (arba harmoniniu ar kt.) indeksų vidurkiu, o kitu atveju — kvadratiniu (antiharmoniniu ar kt.). Iš to išeina, kad statistiškai metodologiniu atžvilgiu geometriniai indeksai, kai kurių statistikos teoretikų ir dar dažniau kai kurių praktikų be atodairos forsuojami, iš tikrųjų, tegali būti tikslingesni už aritmetinius indeksus tik specifinėse sąlygose.

Be to, šalia geometrinių vidurkių kolektyviniams indeksams atstovauti nemažiau teisių tam atstovavimui turi antipodai — kvadratiniai vidurkiai, arba, tam tikromis sąlygomis, harmoniniai, antiharmoniniai ir kt. skaičių eilutės matematiniai vidurkiai.

Statistiškai metodologiniu atžvilgiu vad. matematiniai indeksai statistikos praktikoj yra daug sunkiau pritaikomi ir kartu mažiau prasmingi, kaip matematiniai vidurkiai. Dalykas toks, kad, kaip žinoma, paprastaisiais, svertiniais ar bet kurios rūšies matematiniais vidurkiais būna vaizduojama visuomet griežtai aprėžto skaičiaus dydžių (baigtinė) statistikos eilutė, kur dispersijos linkmė bei pobūdis duoda objektyvinių duomenų ir statistiškai prasmingiausiam matematiniam vidurkiui parinkti. Visai kitokios sąlygos yra kolektyvinius indeksus skaičiuojant. Visų pirma, jau kiekvienas atskirai paimtas individualinių indeksų eilutės skaičius yra griežtai apibrėžtame funkcionaliniame santykyje su baziniais skaičiais. Šių skaičių parinkimas ar pasirinkimas pareina nuo teoretinių bei statistiškai metodologinių samprotavimų ir net nuo skaičiavimo technikos reikalavimų. Šitie gi samprotavimai, kurie nulemia kiekvieną atskirai paimtų individualinių indeksų eilutės narių skaičių, dažnai nulemia ir to skaičiaus lyginamąjį svorį kolektyviniame indekse, — taigi ligi tam tikro laipsnio predestinuoja ir kolektyviniam indeksui imamos skaičių eilutės dispersiją. Tačiau, samprotavimai, kurie nulemia, pav., vienokios ar kitokios bazės pasirinkimą individualiniams indeksams, dažnai nesiduoda statistikos metodais kontroliuojami. O iš to seka, kad bet kurio reiškinio kitimus atstovaujantieji indeksiniai skaičiai statistiškai vienaprasmiai tėra tik to reiškinio atskiriems momentams arba periodams (pav., vienaprasmiai tėra rugių kainų kitimai sausio, vasario ar kt. mėnesiams, bet ne vienaprasmiai rugių ir linų kainos skirtumai tą patį mėnesį). Savaime tie indeksiniai skaičiai dar neduoda jokių garantijų, kad ir gretimųjų b, c, d etc. reiškinių (prekių kainos) skirtumus ar kitimus atstovaujantieji indeksiniai skaičiai (tiems patiems mėnesiams, metams, ar kt.) vienodais kitimais pasireikš ir visą laiką ateityje.

Kolektyviniai indeksai turi reprezentuoti individualinių indeksų kolektyvą vieningai atskiriems laiko periodams arba statistikos stebimiems momentams, kurie gali tęstis be galo. Tačiau, jeigu indeksų kolektyvą norima vaizduoti griežtai matematiniu indeksu, tai kiekvienam vaizduojamam momentui arba periodui matematinis indeksas turėtų atitikti k i e k v i e n o konkretiškai paimtojo momento (arba periodo) skaičių eilutės dispersijos ypatybes. O tai reiškia, kad jeigu, sakysim, sausio mėn. A prekių grupės kainų indeksiniai skaičiai aritmetinį vidurkį turėtų mažesnį už medianą, tai kolektyvinis tos grupės indeksas privalėtų būti kvadratinis, o jeigu vasario mėn. tai pačiai prekių grupei A individualinių indeksų aritmetinis vidurkis būtų didesnis už medianą, tai reikalingas būtų geometrinis indeksas. Tačiau šitoks elgesys būtų tolygus svorio koeficientų keitimui svertinių indeksų sistemoj, ir kolektyvinių indeksų eilutės skaičiai nesiduotų lyginami. Ir nežinotume, pav., ar bet kurios prekių grupės kainų bendrojo indekso rodomieji kitimai pareina nuo tos grupės bendrojo kainolygio krypties pasikeitimo, ar nuo dispersijos pobūdžio pakrypimo.

Statistiškai metodologiniu atžvilgiu indeksavimo formulos taikymas kiekvieno atskirai paimto momento individualinių indeksų eilutės dispersijos ypatybėms arčiausia atitiktų kolektyvinius svertinius indeksus su kintamaisiais svorio koeficientais iš indeksuojamo periodo (ar momento) sulig LXVI formulos:

$$I = \frac{\Sigma p_i q_i}{\Sigma p_o q_i}.$$

Skirtumas tas, kad svorio koeficientai keičiami tik atsižvelgiant vieno ar kito objektyviškai aprėžiamo reiškinio (pav., vienos ar kitos prekės bendrosios tos apyvartos lyginamojo svorio pakitimų). Individualinių gi indeksų eilutės dispersija gali pareiti vien nuo vienos ar keletos prekių kainos žymesnio pakitimo, o ne nuo tos ar kitos prekės lyginamojo svorio tautos ūkyje (apyvartos, gamybos ar vartojimo) pasikeitimo.

Iš to išeina, kad vad. matematiniai indeksai statistiškai metodologiniu atžvilgiu virsta statistiškai nehomogeniniais, t. y. nevienalyčiais dydžiais, — jei taikyti jiems griežtus matema-

tinio analizio dėsnius kiekvieno indeksuojamo momento arba periodo skaičių eilutei, ir pasidaro statistiškai neprasmingi skaičiai, — jei tų dėsnių neatsižvelgti. Tokie indeksai, be to, gali tikrenybę visai iškreipti (statistiškai matodologiniu atžvilgiu falsifikuoti), jei indeksuot pasikliaujant viena kuria matematinių indeksų formula (naudoti vien geometrinius, kvadratinius ar kt. indeksus), visai neatsižvelgiant indeksuojamos skaičių eilutės dispersijos ypatybių.

g) Irving Fischer'io „ideališkoji" indeksavimo formula.

Irving Fischer'is, kuris išanalizavo įvairiausius kolektyvinių indeksų sudarymo metodus, iš 44 išanalizuotųjų formulų siūlo „ideališkąją" indeksavimo formulą, kuri yra sukombinuota iš dviejų svertinių indeksų geometrinio vidurkio. Iš tų dviejų svertinių indeksų vienas gaunamas sulig pastoviais svorio koeficientais iš bazinio periodo (Laspeyres metodas), o kitas sulig kintamais svorio koeficientais iš indeksuojamo periodo (Paasche metodas). Tuo būdu „ideališkoji" indeksavimo formula duos:

$$I_i = \sqrt{\frac{\Sigma p_i q_o}{\Sigma p_o q_o} \cdot \frac{\Sigma p_i q_i}{\Sigma p_o q_i}} \qquad \text{(LXVIII).}$$

Formališkai imant, I_i Fischer'io formula gali būti laikoma tikrai ideališka ypač dėl to, kad, matematinį indeksavimo metodą aprėžus dviejų dydžių vidurkiu, atkrinta reikalas laiko bėgyje matematinio indeksavimo formulą keisti bei taikyti sulig ilgesnės statistinės eilutės dispersijos ypatybių. Statistiškai metodologiniu atžvilgiu ir čia galima prikišti savivalingą g e o m e t r i n i o vidurkio pasirinkimą, kai savaime bet kuris kitas matematinis vidurkis (harmoninis, kvadratinis, antiharmoninis ar kt.) dviem duotiem dydžiam nemažiau gali būti vadinamas charakteringu, kaip geometrinis. Bet ir šitas priekaištas šiose sąlygose yra daugiau formalinis, nes statistikoj gana dažnai tenka kurį nors vieną iš keleto vienodai galimų kriterijų pasirinkti bet kurį pažymį aprėžiant. Svarbiausia, kad šiuo atveju dviejų dydžių geometriniu vidurkiu niekas negali būti suklaidintas, nes dviejų dydžių geometrinis vidurkis

visuomet yra griežtoj pareinamybėj su kitais žinomais matematiniais vidurkiais ir garantuoja visuomet bent kiek mažesnius indekso kitimus laiko bėgyje. Taigi šiuo atžvilgiu I. Fisher'io formula tikrai gali būti laikoma patenkinama.

Statistinės technikos ir praktikos atžvilgiu Fischer'io „ideališkas" indeksas, visų pirma, yra reikalingas žymiai daugiau darbo: nes pirmiau reikalinga dviem skirtingais metodais išskaičiuoti svertinius indeksus ir tik paskui jau ieškoti „ideališkojo" geometrinio indekso, t. y. dviejų svertinių indeksų geometrinio vidurkio. Tačiau svertiniams indeksams reikalingųjų svorio koeficientų problema palieka kaip buvus neišspręsta. Taigi ir I. Fischer'io siūlomo indekso ideališkumas galų gale pareis ne tiek nuo formulos, kiek nuo sugebėjimo statistiškai ideališkus svorio koeficientus surasti. O ir čia dar gali būti abejojama, ar dviem skirtingais metodais gautųjų svertinių indeksų mišinys gali būti laikoma ideališka statistinio vaizdavimo priemone tuo metu, kai statistikos teorija reikalauja kiek galint didesnio statistinių dydžių homogeniškumo.

Skaičiavimo pavyzdžiu galima paimti lentelės 54 a b c duomenis, kur gyvulių kainų svertiniai kolektyviniai indeksai išskaičiuoti skirtingais metodais.

Aplamai, pagal I. Fischer'io formulą išeina:

$$I_{a,i} = \sqrt{\frac{\Sigma p_i q_o}{\Sigma p_o q_o} \cdot \frac{\Sigma p_i q_i}{\Sigma p_o q_i}} :$$

$$I_{a,1} = \sqrt{83{,}7 \,.\, 83{,}6} = \sqrt{6997{,}32} \;(=)\; 83{,}66$$

$$I_{a,2} = \sqrt{115{,}2 \,.\, 114{,}9} = \sqrt{13236{,}48} \;(=)\; 114{,}99$$

$$I_{b,i} = \sqrt{\frac{\Sigma i_i q_o}{\Sigma i_o q_o} \cdot \frac{\Sigma i_i q_i}{\Sigma i_i q_o}} :$$

$$I_{b,1} = \sqrt{95{,}3 \,.\, 96{,}7} = \sqrt{9215{,}51} \;(=)\; 95{,}9$$

$$I_{b,2} = \sqrt{121{,}2 \,.\, 112{,}4} = \sqrt{13622{,}88} \;(=)\; 116{,}6.$$

h) Skirtingais metodais gautųjų kolektyvinių indeksų palyginimas.

Be eilinių lentelių, kur demonstruojamas individualinių ir kolektyvinių indeksų skaičiavimas skirtingais metodais bei to paties metodo skirtingais varijantais, čia patiekiami skirtingais metodais gautieji kolektyviniai indeksai sugretinti visi kartu.

1. Aritmetiniai indeksai:

	I_1	I_2
$I_a = \frac{\Sigma p_i}{\Sigma p_o} 100$	84,0	112,6
$I_b = \frac{\Sigma i_i}{\Sigma i_o} 100$	115,6	104,7

2. Indeksai su pastoviais svorio koeficientais:

$I_a = \frac{\Sigma p_i q_o}{\Sigma p_o q_o} 100$	83,7	115,2
$I_b = \frac{\Sigma i_i q_i}{\Sigma i_o q_o} 100$	95,3	121,2

3. Indeksai su kintamais svorio koeficientais:

$I_a = \frac{\Sigma p_i q_i}{\Sigma p_o q_i} 100$	83,6	114,9
$I_b = \frac{\Sigma i_i q_i}{\Sigma i_o q_i} 100$	96,7	112,4

4. Irving Fischer'io indeksai:

$I_a = \sqrt{\frac{\Sigma p_1 q_o}{\Sigma p_o q_o} \cdot \frac{\Sigma p_i q_i}{\Sigma p_o q_o}}$	83,66	114,99
$I_b = \sqrt{\frac{\Sigma i_i q_o}{\Sigma i_o q_o} \cdot \frac{\Sigma i_i q_i}{\Sigma i_o q_i}}$	95,9	116,6

Pirmiausia čia krinta į akis ta aplinkybė, kad kiekvienas iš nagrinėtųjų metodų duoda skirtingus indeksus sulig tuo, ar

jie buvo skaičiuoti betarpiškai iš originalinių st. eilučių, ar iš jas atstovaujančių individualinių indeksų; be to, paprastasis aritmetinis metodas davė ne tik skirtingus kolektyvinius indeksus, bet ir skirtingą tendenciją (tiems patiems dydžiams kolektyvinis indeksas I_2 iš individualinių indeksų yra mažesnis už anksčiau buvusį I_1, o kolektyvinis indeksas betarpiškai iš originalinių yra didesnis už I_1).

Svertiniai kolektyviniai indeksai savitarpio tendencijos duotiesiems skaičiams nėra pakeitę, tačiau kolektyvinių indeksų skirtumai yra didesni ir ryškesni, kai skirtingi (pastovūs arba kintamieji) svorio koeficientai taikomi individualiniams indeksams.

Irving Fischer'io vad. ideališkieji indeksai, atrodo, turi didesnės prasmės tik tuo atveju, kai skaičiuojami individualinių indeksų kolektyvams, o ne originalinių skaičių eilutėms.

Aplamai, galima sakyti, kad kolektyviniams indeksams svarbiausias dalykas yra statistinių duomenų tikrumas bei indeksų skaičiavimo metodo ir bazės vienodumas. Ir juo didesnis kolektyviniam indeksui naudojamų st. eilučių (ar individualinių indeksų) skaičius, tuo mažesnės reikšmės turi indeksų sudarymo metodas. O tokiais atvejais įgauna didesnės prasmės indeksų skaičiavimo technikos patogumai, t. y. panaudojami paprastesnieji metodai, kurie mažiau reikalingi darbo (paprastieji aritmetiniai indeksai arba svertiniai indeksai su pastoviais svorio koeficientais).

Literatūra ir papildomos žinios. Apie indeksus yra visuose naujesniųjų laikų statistikos vadovėliuose, be to, 1) P. Flaskämper, Theorie der Indexzahlen, Berlin, 1928, 200 p. 2) Irving Fischer, The Making of Index Numbers, Boston and New York 1923.

Kvadratai ir kūbai

kvadratinės ir kūbinės šaknys, atverstieji (reciprokiniai) skaičiai ir logaritmai natūriniams skaičiams nuo 1 ligi 100.

n	n^2	n^3	$\sqrt{n}$	$\sqrt[3]{n}$	$\frac{1}{n}$	log n
0	0	0	0,0000	0,0000	∞	$-\infty$
1	1	1	1,0000	1,0000	1,00000	0,0000
2	4	8	1,4142	1,2599	0,50000	0,3010
3	9	27	1,7321	1,4422	0,33333	0,4771
4	16	64	2,0000	1,5874	0,25000	0,6021
5	25	125	2,2361	1,7100	0,20000	0,6990
6	36	216	2,4495	1,8171	0,16667	0,7781
7	49	343	2,6458	1,9129	0,14286	0,8451
8	64	512	2,8284	2,0000	0,12500	0,9031
9	81	729	3,0000	2,0801	0,11111	0,9542
10	100	1000	3,1623	2,1544	0,10000	1,0000
11	121	1331	3,3166	2,2240	0,09091	1,0414
12	144	1728	3,4641	2,2894	0,08333	1,0792
13	169	2197	3,6056	2,3513	0,07692	1,1139
14	196	2744	3,7417	2,4101	0,07143	1,1461
15	225	3375	3,8730	2,4662	0,06667	1,1761
16	256	4096	4,0000	2,5198	0,06250	1,2041
17	289	4913	4,1231	2,5713	0,05882	1,2304
18	324	5832	4,2426	2,6207	0,05556	1,2553
19	361	6859	4,3589	2,6684	0,05263	1,2788
20	400	8000	4,4721	2,7144	0,05000	1,3010
21	441	9261	4,5826	2,7589	0,04762	1,3222
22	484	10648	4,6904	2,8020	0,04545	1,3424
23	529	12167	4,7958	2,8439	0,04348	1,3617
24	576	13824	4,8990	2,8845	0,04167	1,3802
25	625	15625	5,0000	2,9240	0,04000	1,3979
26	676	17576	5,0990	2,9625	0,03846	1,4150
27	729	19683	5,1962	3,0000	0,03704	1,4314
28	784	21952	5,2915	3,0366	0,03571	1,4472
29	841	24389	5,3852	3,0723	0,03448	1,4624
30	900	27000	5,4772	3,1072	0,03333	1,4771
31	961	29791	5,5678	3,1414	0,03226	1,4914
32	1024	32768	5,6569	3,1748	0,03125	1,5052
33	1089	35937	5,7446	3,2075	0,03030	1,5185
34	1156	39304	5,8310	3,2396	0,02941	1,5315
35	1225	42875	5,9161	3,2711	0,02857	1,5441
36	1296	46656	6,0000	3,3019	0,02778	1,5563
37	1369	50653	6,0828	3,3322	0,02703	1,5682
38	1444	54872	6,1644	3,3620	0,02632	1,5798
39	1521	59319	6,2450	3,3912	0,02564	1,5911
40	1600	64000	6,3246	3,4200	0,02500	1,6021
41	1681	68921	6,4031	3,4482	0,02439	1,6128
42	1764	74088	6,4807	3,4760	0,02381	1,6232
43	1849	79507	6,5574	3,5034	0,02326	1,6335
44	1936	85184	6,6332	3,5303	0,02273	1,6435
45	2025	91125	6,7082	3,5569	0,02222	1,6532
46	2116	97336	6,7823	3,5830	0,02174	1,6628
47	2209	103823	6,8557	3,6088	0,02128	1,6721
48	2304	110592	6,9282	3,6342	0,02083	1,6812
49	2401	117649	7,0000	3,6593	0,02041	1,6902
50	2500	125000	7,0711	3,6840	0,02000	1,6990

n	n^2	n^3	$\sqrt{n}$	$\sqrt[3]{n}$	$\frac{1}{n}$	log n
51	2601	132651	7,1414	3,7084	0,01961	1,7076
52	2704	140608	7,2111	3,7325	0,01923	1,7160
53	2809	148877	7,2801	3,7563	0,01887	1,7243
54	2916	157464	7,3485	3,7798	0,01852	1,7324
55	3025	166375	7,4162	3,8030	0,01818	1,7404
56	3136	175616	7,4833	3,8259	0,01786	1,7482
57	3249	185193	7,5498	3,8485	0,01754	1,7559
58	3364	195112	7,6158	3,8709	0,01724	1,7634
59	3481	205379	7,6811	3,8930	0,01695	1,7709
60	3600	216000	7,7460	3,9149	0,01667	1,7782
61	3721	226981	7,8102	3,9365	0,01639	1,7853
62	3844	238328	7,8740	3,9579	0,01613	1,7924
63	3969	250047	7,9373	3,9791	0,01587	1,7993
64	4096	262144	8,0000	4,0000	0,01563	1,8062
65	4225	274625	8.0623	4,0207	0,01538	1,8129
66	4356	287496	8,1240	4,0412	0,01515	1,8195
67	4489	300763	8,1854	4,0615	0,01493	1,8261
68	4624	314432	8,2462	4,0817	0,01471	1,8325
69	4761	328509	8,3066	4,1016	0,01449	1,8388
70	4900	343000	8,3666	4,1213	0,01429	1,8451
71	5041	357911	8,4261	4,1408	0,01408	1,8513
72	5184	373248	8,4853	4,1602	0,01389	1,8573
73	5329	389017	8,5440	4,1793	0,01370	1,8633
74	5476	405224	8,6023	4,1983	0,01351	1,8692
75	5625	421875	8,6603	4,2172	0,01333	1,8751
76	5776	438976	8,7178	4,2358	0,01316	1,8808
77	5929	456533	8,7750	4,2543	0,01299	1,8865
78	6084	474552	8,8318	4,2727	0,01282	1,8921
79	6241	493039	8,8882	4,2908	0,01266	1,8976
80	6400	512000	8,9443	4,3089	0,01250	1,9031
81	6561	531441	9,0000	4,3267	0,01235	1,9085
82	6724	551368	9,0554	4,3445	0,01220	1,9138
83	6889	571787	9,1104	4,3621	0,01205	1,9191
84	7056	592704	9,1652	4,3795	0,01190	1,9243
85	7225	614125	9,2195	4,3968	0,01176	1,9294
86	7396	636056	9,2736	4,4140	0,01163	1,9345
87	7569	658503	9,3274	4,4310	0,01149	1,9395
88	7744	681472	9,3808	4,4480	0,01136	1,9445
89	7921	704969	9,4340	4,4647	0,01124	1,9494
90	8100	729000	9,4868	4,4814	0,01111	1,9542
91	8281	753571	9,5394	4,4979	0,01099	1,9590
92	8464	778688	9,5917	4,5144	0,01087	1,9638
93	8649	804357	9,6437	4,5307	0,01075	1,9685
94	8836	830584	9,6954	4,5468	0,01064	1,9731
95	9025	857375	9,7468	4,5629	0,01053	1,9777
96	9216	884736	9,7980	4,5789	0,01042	1,9823
97	9409	912673	9,8489	4,5947	0,01031	1,9868
98	9604	941192	9,8995	4,6104	0,01020	1,9912
99	9801	970299	9,9499	4,6261	0,01010	1,9956
100	10000	1000000	10,0000	4,6416	0,01000	2,0000

Kvadratinės ir kūbinės šaknys

dešimt. trupmenoms nuo 0,01 ligi 0,9 ir paprastosioms trupmenoms nuo $\frac{1}{12}$ ligi $\frac{7}{8}$.

n	$\sqrt{n}$	$\sqrt[3]{n}$	n	$\sqrt{n}$	$\sqrt[3]{n}$	n	$\sqrt{n}$	$\sqrt[3]{n}$	n	$\sqrt{n}$	$\sqrt[3]{n}$
0,01	0,100	0,215	0,25	0,500	0,630	$^1/_4$	0,500	0,630	$^3/_8$	0,612	0,721
0,02	0,141	0,271	0,3	0,548	0,669	$^3/_4$	0,866	0,909	$^5/_8$	0,791	0,855
0,03	0,173	0,311	0,4	0,632	0,737	$^1/_6$	0,408	0,550	$^7/_8$	0,935	0,956
0,04	0,200	0,342	0,5	0,707	0,794	$^5/_6$	0,913	0,941	$^1/_9$	0,333	0,481
0,05	0,224	0,368	0,6	0,775	0,843	$^1/_7$	0,378	0,523	$^2/_9$	0,471	0,606
0,06	0,245	0,391	0,7	0,837	0,888	$^2/_7$	0,535	0,659	$^4/_9$	0,667	0,763
0,07	0,265	0,412	0,75	0,867	0,909	$^3/_7$	0,655	0,754	$^5/_9$	0,745	0,822
0,08	0,283	0,431	0,8	0,894	0,928	$^4/_7$	0,756	0,830	$^7/_9$	0,882	0,920
0,09	0,300	0,448	0,9	0,949	0,965	$^5/_7$	0,845	0,894	$^1/_{12}$	0,289	0,437
0,1	0,316	0,464	$^1/_3$	0,577	0,693	$^6/_7$	0,926	0,950	$^5/_{12}$	0,645	0,747
0,2	0,447	0,585	$^2/_3$	0,817	0,874	$^1/_8$	0,354	0,500	$^7/_{12}$	0,764	0,836

Tikimybių lentelė.

Tikimybės, kurios atitinka t, t. y. vidutinės palaidos ū daugiklį (event. vidurinio kvadratingojo nukrypimo σ daugiklį).

t	Tikimybės	t	Tikimybės	t	Tikimybės	t	Tikimybės
0,0	0,00000	1,0	0,68268	2,0	0,95450	3,0	0,99730
0,1	0,07966	1,1	0,72866	2,1	0,96428	3,1	0,99806
0,2	0,15852	1,2	0,76986	2,2	0,97220	3,2	0,99862
0,3	0,23582	1,3	0,80640	2,3	0,97856	3,3	0,99904
0,4	0,31078	1,4	0,83848	2,4	0,98360	3,4	0,99932
0,5	0,38292	1,5	0,86638	2,5	0,98758	3,5	0,99954
0,6	0,45150	1,6	0,89040	2,6	0,99068	3,6	0,99968
0,7	0,51608	1,7	0,91086	2,7	0,99306	3,7	0,99978
0,8	0,57628	1,8	0,92814	2,8	0,99488	3,8	0,99986
0,9	0,63188	1,9	0,94256	2,9	0,99626	3,9	0,99990
1,0	0,68268	2,0	0,95450	3,0	0,99730	4,0	0,99994

Funkcijos X_i[1])

N nelyginis skaičius.

x	X_2	X_3	X_4	X_5
	N=3		$X_1(1) = 1{,}2247449$	
0	−1,41421			
1	0,70711			
Σ	−0,70710			
	N=5		$X_1(1) = 0{,}7071068$	
0	−1,19523	0,00000	1,60357	
1	−0,59761	−1,41421	−1,06905	
2	1,19523	0,70711	0,26726	
Σ	−0,59761	−0,70710	0,80178	
	N=7		$X_1(1) = 0{,}5000000$	
0	−1,15470	0,00000	1,27920	0,00000
1	−0,86603	−1,08012	0,21320	1,44338
2	0,00000	−1,08012	−1,49241	−1,15470
3	1,44338	1,08012	0,63960	0,28868
Σ	−0,57735	−1,08012	0,63959	0,57736
	N=9		$X_1(1) = 0{,}3872983$	
0	−1,13961	0,00000	1,20687	0,00000
1	−0,96866	−0,85812	0,60344	1,24808
2	−0,45584	−1,23950	−0,73753	0,55470
3	0,39886	−0,66742	−1,40802	−1,52543
4	1,59545	1,33485	0,93868	0,55470
Σ	−0,56980	−1,43019	0,60344	0,83205

N lyginis skaičius.

x	X_2	X_3	X_4	X_5
	N=4			$X_1(1) = 0{,}4472136$
1	−1,00000	−1,34164		
3	1,00000	0,44721		
Σ	0	−0,89443		
	N=6			$X_1(1) = 0{,}2927700$
1	−1,06905	−0,73030	0,92582	1,54303
3	−0,26726	−1,27802	−1,38873	−0,77152
5	1,33631	0,91287	0,46291	0,15430
Σ	0	−1,09545	0	0,92581
	N=8			$X_1(1) = 0{,}2182179$
1	−1,09109	−0,52223	1,02564	0,90784
3	−0,65465	−1,21854	−0,34188	1,02889
5	0,21822	−0,87039	−1,48149	−1,39202
7	1,52752	1,21854	0,79772	0,42366
Σ	0	−1,39262	−1	0,96837
	N=10			$X_1(1) = 0{,}1740777$
1	−1,10096	−0,40967	1,06436	0,67937
3	−0,82572	−1,05832	0,17739	1,24550
5	−0,27524	−1,19488	−1,00523	0,11323
7	0,55048	−0,47795	−1,30089	−1,58519
9	1,65145	1,43386	1,06436	0,67937
Σ	1	−1,70696	−1	1,13228

[1] Svarbu įsidėmėti, kad $X_1(2)=2X_1(1)$, $X_1(3)=3X_1(1)$, ir t. t. Taigi pirmaeilė linkmė gaunama šiuo būdu:

kai N nelyginis sk.: a_0, $a_0 \pm a_1X_1(1)$, $a_0 \pm 2.a_1X_1(1)$, $a_0 \pm 3.a_1X_1(1), \ldots, a_0 \pm \frac{N-1}{2} a_1X_1(1)$;

kai N lyginis sk.: $a_0 \pm a_1X_1(1)$, $a_0 \pm 3.a_1X_1(1)$, $a_0 \pm 5.a_1X_1(1), \ldots, a_0 \pm (N-1)a_1X_1(1)$.

X_i funkcijos ir k_i bei l_i faktorių skaičiai paimti iš Lorenc'o „Trend".

N nelyginis skaičius.

x	X_2	X_3	X_4	X_5
	N=11		$X_1(1) = 0{,}3162278$	
0	—1,13228	0,00000	1,17670	0,00000
1	—1,01905	—0,70892	0,78446	1,06217
2	—0,67937	—1,16465	—0,19612	1,06217
3	—0,11323	—1,11401	—1,17670	—0,26554
4	0,67937	—0,30382	—1,17670	—1,59326
5	1,69842	1,51911	1,17670	0,79663
Σ	—0,56614	—1,77229	0,58834	1,06217
	N=13		$X_1(1) = 0{,}?672612$	
0	—1,12815	0,00000	1,16086	0,00000
1	—1,04757	—0,60302	0,88446	0,91670
2	—0,80582	—1,05529	0,15202	1,19171
3	—0,40291	—1,20605	—0,74627	0,50418
4	0,16116	—0,90453	—1,32670	—0,82503
5	0,88640	0,00000	—0,91210	—1,51255
6	1,77281	1,65831	1,36815	1,00837
Σ	—0,56408	—2,11058	0,58042	1,28338
	N=15		$X_1(1) = 0{,}2314550$	
0	—1,12560	0,00000	1,15142	0,00000
1	—1,06530	—0,52430	0,94581	0,80367
2	—0,88440	—0,95150	0,38228	1,19062
3	—0,58290	—1,18452	—0,37924	0,89415
4	—0,16080	—1,12627	—1,07222	—0,05239
5	0,38190	—0,67964	—1,32352	—1,16561
6	1,04520	0,25244	—0,65339	—1,36207
7	1,82909	1,76707	1,52457	1,19181
Σ	—0,56281	—2,44672	0,57571	1,50018

N lyginis skaičius.

x	X_2	X_3	X_4	X_5
	N=12		$X_1(1) = 0{,}1448414$	
1	—1,10624	—0,33796	1,08389	0,54924
3	—0,91660	—0,91733	0,46453	1,20832
5	—0,53732	—1,20701	—0,50324	0,79639
7	0,03161	—1,01389	—1,27745	—0,57670
9	0,79017	—0,14484	—1,04518	—1,56532
11	1,73838	1,59326	1,27745	0,90624
Σ	0	—2,02777	0	1,31817
	N=14		$X_1(1) = 0{,}1240347$	
1	—1,10940	—0,28797	1,09522	0,46297
2	—0,97073	—0,80393	0,63888	1,11883
5	—0,69338	—1,13990	—0,13183	1,07254
7	—0,27735	—1,17589	—0,93297	0,21605
9	0,27735	—0,79193	—1,33860	—1,01852
11	0,97073	0,13199	—0,78085	—1,44291
13	1,80278	1,71584	1,45015	1,10340
Σ	0	—2,35179	0	1,51236
	N=16		$X_1(1) = 0{,}1084652$	
1	—1,11144	—0,25103	1,10240	0,40094
3	—1,00559	—0,71324	0,75243	1,02162
5	—0,79388	—1,05591	0,13415	1,16718
7	—0,47633	—1,19936	—0,58911	0,68605
9	—0,05293	—1,06388	—1,17240	—0,29402
11	0,47633	—0,56979	—1,28905	—1,27410
13	1,11144	0,36260	—0,53079	—1,27410
15	1,85240	1,81298	1,59236	1,27410
Σ	0	—2,67763	—	1,71067

N nelyginis skaičius.

x	X_2	X_3	X_4	X_5
	N=17		$X_1(1) = 0{,}2041241$	
0	—1,12390	0,00000	1,14531	0,00000
1	—1,07707	—0,46359	0,98624	0,71435
2	—0,93659	—0,86095	0,54084	1,14295
3	—0,70244	—1,12585	—0,09544	1,07801
4	—0,37463	—1,19208	—0,76354	0,46757
5	0,04683	—0,99340	—1,24076	—0,50654
6	0,56195	—0,46359	—1,24076	—1,35076
7	1,17073	0,46359	—0,41359	—1,18192
8	1,87317	1,85434	1,65434	1,35076
Σ	—0,56195	—2,78153	0,57264	1,71442
	N=19		$X_1(1) = 0{,}1825742$	
0	—1,12272	0,00000	1,14112	0,00000
1	—1,08530	—0,41539	1,01433	0,64235
2	—0,97303	—0,78358	0,65413	1,08032
3	—0,78591	—1,05736	0,12103	1,15332
4	—0,52394	—1,18953	—0,48411	0,78834
5	—0,18712	—1,13288	—1,02009	0,04380
6	0,22454	—0,84022	—1,30537	—0,84674
7	0,71106	—0,26434	—1,11807	—1,43069
8	1,27242	0,64197	—0,19595	—0,99273
9	1,90863	1,92590	1,76355	1,48909
Σ	—0,56137	—3,11543	0,57057	1,92706

N lyginis skaičius.

x	X_2	X_3	X_4	X_5
	N=18		$X_1(1) = 0{,}0963739$	
1	—1,11283	—0,22257	1,10725	0,35396
3	—1,02937	—0,63988	0,83044	0,93800
5	—0,86244	—0,97373	0,32714	1,17933
7	—0,61206	—1,16847	—0,30198	0,94604
9	—0,27821	—1,16847	—0,90593	0,25099
11	0,13910	—0,91808	—1,28341	—0,69022
13	0,63988	—0,36167	—1,18275	—1,40137
15	1,22411	0,55641	—0,30198	—1,08763
17	1,89181	1,89181	1,71121	1,42228
Σ	— 1	—3,00465	— 1	1,91138
	N=20		$X_1(1) = 0{,}0867110$	
1	—1,11382	—0,19995	1,11068	0,31704
3	—1,04632	—0,57964	0,88630	0,86146
5	—0,91131	—0,89875	0,47026	1,15369
7	—0,70880	—1,11687	—0,07199	1,08163
9	—0,43878	—1,19362	—0,64229	0,61728
11	—0,10126	—1,08860	—1,10975	—0,14972
10	0,30377	—0,76141	—1,31076	—0,97835
15	0,77630	—0,17167	—1,04898	—1,44271
17	1,31634	0,72102	—0,09536	—0,89829
19	1,92388	1,95705	1,81187	1,55160
Σ	0	—3,33244	— 2	2,11363

N nelyginis skaičius. N lyginis skaičius.

x	X_2	X_3	X_4	X_5	x	X_2	X_3	X_4	X_5
	N=21		$X_1(1) = 0{,}1651446$			N=22			$X_1(1) = 0{,}0788110$
0	−1,12186	0,00000	1,13811	0,00000	1	−1,11456	−0,18153	1,11321	0,28722
1	−1,09127	−0,37622	1,03465	0,58325	3	−1,05883	−0,52945	0,92767	0,79464
2	−0,99948	−0,71761	0,73767	1,01529	5	−0,94737	−0,83200	0,57880	1,11132
3	−0,84650	−0,98933	0,28740	1,17107	5	−0,78019	−1,05891	0,11100	1,14446
4	−0,63232	−1,15654	−0,24908	0,97790	9	−0,55728	−1,17992	−0,40913	0,85282
5	−0,35696	−1,18440	−0,77790	0,44159	11	−0,27864	−1,16480	−0,89278	0,26734
6	−0,02040	−1,03810	−1,17835	−0,32735	13	0,05573	−0,98327	−1,22897	−0,48827
7	0,37735	−0,68277	−1,30289	−1,08757	15	0,44582	−0,60509	−1,28447	−1,17686
8	0,83630	−0,08361	−0,97717	−1,44068	17	0,89165	0,00000	−0,90389	−1,42726
9	1,35644	0,79425	0,00000	−0,80508	19	1,39320	0,86225	0,09039	−0,71363
10	1,93777	1,98562	1,85662	1,61017	21	1,95047	2,01192	1,89816	1,66513
Σ	−0,56093	−3,44871	0,56906	2,13859	Σ	0	−3,66080	— 1	2,31691
	N=23		$X_1(1) = 0{,}1507557$			N=24			$X_1(1) = 0{,}0722315$
0	−1,12122	0,00000	1,13588	0,00000	1	−1,11511	−0,16623	1,11511	0,26260
1	−1,09574	−0,34378	1,04983	0,53394	3	−1,06833	−0,48707	0,95915	0,73637
2	−1,01929	−0,66111	0,80094	0,95287	5	−0,97475	−0,77303	0,66283	1,06250
3	−0,89188	−0,92555	0,41702	1,15823	7	−0,83439	−1,00088	0,25733	1,16460
4	−0,71351	−1,11066	−0,05560	1,08430	9	−0,64723	−1,14734	−0,21055	0,99933
5	−0,48416	−1,18999	−0,55205	0,71465	11	−0,41329	−1,18919	−0,67843	0,56963
6	−0,20386	−1,13710	−0,98893	0,09857	13	−0,13257	−1,10317	−1,06833	−0,06207
7	0,12741	−0,92555	−1,26430	−0,63251	15	0,19495	−0,86603	−1,28667	−0,76061
8	0,50965	−0,52889	−1,25768	−1,24859	17	0,56925	−0,45452	−1,22429	−1,30490
9	0,94285	0,07933	−0,83007	−1,40466	19	0,99035	0,15461	−0,75641	1,37468
10	1,42701	0,92555	0,17607	−0,62429	21	1,45823	0,98460	0,25733	−0,53731
11	1,96214	2,03621	1,93682	1,71681	23	1,97289	2,05871	1,97289	1,76545
Σ	−0,56060	−3,78154	0,56793	2,34932	Σ	0	−3,98954	— 4	2,52091

Faktorius k_j, kai N yra nelyginis skaičius.

Faktorius l_j, kai N — nelyginis skaičius.

N	k_0	k_1	k_2	k_3	k_4	k_5	N	l_3	l_4	l_5
3	1,41421	408,248	707,107	—	—	—	3	—	—	—
5	1 19523	141,421	119,523	353,553	4183,30	—	5	2,00000	0,447214	—
7	1,15470	71,4286	41,2393	71,4286	577,350	3000,00	7	2,16025	0,369274	0,577350
9	1,13961	43,0331	18,9934	23,9073	147,727	430,331	9	2,21565	0,353009	0,644503
11	1,13228	28,7480	10,2934	10,2671	51,'671	107,805	11	2,24179	0,346410	0,671775
13	1,12815	20,5586	6,19864	5,13964	21,6952	35.9775	13	2,25630	0,342997	0,685994
15	1,12560	15,4303	4,01999	2,85747	10,4222	14,4016	15	2,26522	0,340981	0,694448
17	1,12390	12,0073	2,75466	1,71533	5,50933	6,54944	17	2,27110	0,339684	0,699913
19	1,12272	9,60917	1,96969	1,09195	3,13359	3,27585	19	2,27519	0,338796	0,703662
21	1,12187	7,86403	1,45697	0,728151	1,88866	1,76437	21	2,27814	0,338161	0,706351
23	1,12122	6,55459	1,10793	0,504200	1,19315	1,00840	23	2,28035	0,337691	0,708346

Faktorius k_j, kai N yra lyginis skaičius.

Faktorius l_j, kai N — lyginis skaičius.

N	k_0	k_1	k_2	k_3	k_4	k_5	N	l_3	l_4	l_5
4	1,00000	111,803	500 000	372,678	—	—	4	3,00000	—	—
6	1,06905	48,7950	133,631	60,9938	3741,66	3074,09	6	2,49444	0,288675	1,054090
8	1,09109	27,2772	54,5545	18,1848	636,469	381,881	8	2,39317	0,313340	0,832050
10	1,10096	17,4078	27,5241	7,25324	183,494	87,0388	10	2,35339	0,322252	0,780532
12	1,10624	12,0701	15,8035	3,44860	69,1403	27,1578	12	2,33333	0,326599	0,758395
14	1,10940	8,85962	9,90536	1,84575	30,8167	10,3362	14	2,32172	0,329073	0,746509
16	1,11144	6,77908	6,61570	1,07604	15,4366	4,51938	16	2,31436	0,330623	0,739296
18	1,11283	5,35410	4,63679	0,669263	8,43053	2,19032	18	2,30940	0,331662	0,734559
20	1,11382	4 33555	3,37522	0,437934	4,92220	1,14958	20	2,30589	0,332394	0,731267
22	1,11456	3,58232	2,53308	0,298527	3,03104	0,642981	22	2,30332	0,332930	0,728882
24	1,11511	3,00965	1,94950	0,210465	1,94950	0,378837	24	2,30137	0,333333	0,727096

I. VARDŲ RODYKLĖ.

(Skaičiai rodo puslapius.)

Achenwall 19—21, 22.
Anderson O. 213.

Bernoulli J. 29.
Bertillon 46.
Bortkiewicz 31
Bowley 54
Buesching 20

Conring 19, 22

Čebyšev 213, 352
Četverikov 347

Ellis H. 12

Fischer Irving 361, 401, 402, 404
Fischer A. 213
Flaskämper 361
Florance Sergant 200

Galton F. 107
Graunt J. 22, 23, 26, 30

Harper F. H. 200
Hull C. H. 22

James 5
Jastremski 286
Julin 305, 307, 308
Jurskis A. F. 200

King 307, 308

Laspeyres 389—392, 401
Lexis 31
Lorenz 347, 349, 352
Lueder 21
Mayr, Georg von 17, 18, 20, 31, 53, 290
Markov 213
Mueller J. 53, 290

Paasche 390, 391, 401
Pearson 308, 309
Persons 379
Petty W. 22, 24, 25, 26, 30

Quetelet 8, 25—30, 46, 180, 259
Romanowsky 21
Sauerbeck 392, 393
Schloezer 20, 21
Schott 2
Seckendorf 19
Soetbeer 389
Stirling 248, 249, 250
Suessmilch 7, 23, 25, 26, 29, 238

Wagemann 141
Westergaard 141, 240, 241, 247, 248
Winkler 153, 154

Žižek Fr. 54, 290

II. ŽODŽIŲ IR DALYKŲ RODYKLĖ.

Anketa 78
Antgalvis 109
Antrinė statistika 67
Aprašomoji statistika 31
Aritmetiškas trikampis 210, 213
Aritmetinis vidurkis, ž. vidurkis
Asimetrijos koeficientas 196, 201, 217, 218, kvartilinis 202, 203, 204, aritmetinis 206

Bandomasis surašymas 89
Bendrosios Statistikos Departamentas 35
Bendrieji lyginiai, ž. statistiniai lyginiai
Binomo koeficientai, ž. dispersija
Breviarium Augusti 17
Breviarium fiscalium 13
Brutto lyginiai, ž. statistiniai lyginiai
Brūkšnelis 112

Centralinis Statistikos Biuras 36—39, 43—46
Centralizacija (statistikos darbo organizacijoj) 33, 36, 43
Cenzoriai (Romoj) 9

Čebyševo teorema 213

Daliniai, ž. Statistiniai daliniai
Dalinis stebėjimo metodas, ž. Statistinio stebėjimo metodai ir sistemos
Decentralizacija 33
Deklaracija 57, 66
Dekada 63
Demografijos kongresai 46
Dėsningumai, ž. Statistiniai dėsningumai
Diagrama 123, 124
Didžiųjų skaičių dėsnis 8, 27, 29, 54, 55, 79, 103, 239, 240, 270
Dieviškoji santvarka 23, 29
Dispersija 64, 80, 101, 107, 109, 189, 192 sek., 245; p a g r i n d i n i a i t i p a i: 194 sek., normali 194, 208 sek., 211, 212, 215, 217, 220, 261; Pareto kreivė 195; *I* dispersija 195; *U* dispersija 195; binomo koeficientų 210, 213, 245—249; vidutinė arba aritmetinė 205; trijų sigmų taisyklė 101, 212, 213, 220, 243, 261, 262; Čebyševo teorema 213; vidutins aritmetinis nukrypimas 205, 206; vidurinis kvadratingasis nukrypimas 211 sek., 220, 251, 252, 253, 262, jam skaičiuoti metodai 225 sek., jo patikimoji paklaida 263
Dispersijos koeficientas 196, 203 sek., 218; kvartilinis 203, 204; vidutinis aritmetinis 204—206; išlygintoj skaičių eilutėj 358 sek.
Domesday Book 12—15
Dominantas 157, 176 sek., 189, 250
Dvigubi skaičiavimai 60

Ekstrapoliacija 303

Galtono metodas 107, 108
Gimimų lyginiai 140, 141, 143, 145, 146
Gyventojų surašinėjimai senovėje: Romoj 9, 10; žydų 10; Graikijoj 10; Egipte 10; Kinijoj 11; Indijoj 11; Japonijoj 11; viduramžiais: Anglijoj 12; Italijoj 13; Vokietijoj 13; Lietuvoj 14 sek.; naujaisiais laikais 15, 16: Anglijoj 16; Lietu-

voj 16, 35; Prancūzijoj 16; Sicilijoj 15; USA 16; Vokietijoj 16
Grafinis vaizdavimas statistikoj 123 sek., stovio statistikoj 123, 124; judėjimo statistikoj 124 sek.; diagrama 123, 124; kvadragrama 123; kartograma 124; liniagrama 125; kreivės 125; mastelio parinkimas 125 sek.; logaritminis mastelis 126
Grynieji lyginiai, ž. Statistiniai lyginiai
Gubernijos statistikos komitetas 34

Higienos kongresai 46
Hipotetiniai dydžiai 321 sek.

Indeksai, esmė 360, 361; bazė 361; pagrindinė klasifikacija 361; su kintama baze: 361, 367 sek., kumuliatyviniai 367, 369; su pastovia baze: 361, 362 sek., normalieji 367, specifiniai 367, pastovi bazė: aritmetinis vidurkis 363, medianas 364, dominantas 364, 365, „norminiai" dydžiai 365, 366, kraštutiniai dydžiai 366, specialiai parinkti dydžiai 366, 367; periodiniai bei sezoniniai: esmė 369, bazės parinkimas 370, paprastųjų periodų indeksai 370, apréžtųjų periodų indeksai 370—372; normalūs periodiniai 370, 373, jų skaičiavimo metodai: periodinių vidurkių metodas 374—376, periodiškai kolektyvinių indeksų metodas 376—379, Persons'o metodas 379—381, įvairių metodų palyginimas 382—384; individualiniai 361, 384; kolektyviniai: esmė 384, 385, sudarymo sąlygos 385, sudarymo metodai: paprastieji aritmetiniai 385, 386, svertiniai 385, 386, svorio koeficientų parinkimo sąlygos 386—389, svorio koeficientų apibrėžimas 392—397, svėrimo sistemos 389—392, Laspeyres metodas 390, 391, 401, Paasche metodas 390, 401, C. S. Biuro 386, Lietuvos Banko 396, 397, Economist'o 365, 395, 396, Statist'o 365, 395, Sauerbeck'o 365, 392, 395; matematiniai 361, 398 sek., jų pritaikymas praktikai 399—401, Fischer'io formula 401, 402
Individualinis stebėjimo metodas 73 sek., 86
Intervalas 102—107, 113, 114, 117
Interpoliacija, esmė 290—292; tiesinė arba aritmetinė 292—294; grafiškoji 292, 294; geometrinė 294 sek.; vidurkiais 298—299; linkmės dalinių 300—303; Newton'o - Gauss'o formula 301
Inventorius (dvarų) 13
Išlyginimai, esmė 326—328, 358; mažiausiųjų kvadratų metodas 329 sek.; aukštesnės eilės parabolių lygtimis 337, 338; sumavimų metodas 347 sek.; kintamaisiais vidurkiais 338 sek.; kintamųjų parametrų metodas 341 sek.
Ištisinis stebėjimo metodas 71, 81, 82

„Jungas" 14
Jungtuvių lyginiai 141, 143, 145

Kainų žirklės 128, 129, 130
„Kariuomenės sąrašai" (D. K. Lietuvoje) 14
Kartograma 124
Kintamieji vidurkiai 338 sek.; svertiniai 340, 341
Kintamųjų parametrų metodas 341 sek.
Klaipėdos krašto statistikos biuras 39 sek.
Klasifikacija 33, 58; tikslas 95; sistemos 109 sek.; planas 86; esminė 96; laikinė 98; plotinė 96, 97, 98; tikslinė 99; kiekybinė 97, 98, 101 sek.; 113 sek.; kokybinė 97, 98 sek.; intervalas 102—107, 113, 114, 117; statistiškai matematinė 103, 116, 117; pirminė kiekybinė 104 sek., 107;

aukštesniojo laipsnio kiekybinė 106 sek.; Galtono metodas 107, 108, 118; kumuliatyvinis metodas 108, 109, 119; skiriamasis metodas 113, 114; jungiamasis metodas 113, 114; klaidingos klasifikacijos pavyzdžiai 119 sek.
Kolektyvinis stebėjimo metodas 74 sek., 86
Konjunktūros institutai 47, 48
Koreliacija, esmė 304, 305; formalinė pareinamybė 305, 320, 321; kauzalinė pareinamybė 305, 320, 321; koeficientas: 305, 319, 320; Julin formula 305—308
King'o formula 307, 308; Pearson'o formula 308 sek.; skaičiavimas sulig skirtumais nuo vidurkio 309, sulig pirmaisiais skirtumais 309, 313; patikimoji paklaida 312; kolektyvinis 315 sek.; linkmės kitimų hipotetinių dydžių skaičiavimas 321 sek.
Kortelių sistema 58, 73, 74, 86
Kritiškasis momentas 58, 60—62, 64, 65
Kumuliatyvinis metodas, ž. Klasifikacija
Kvadragrama 123
Kvadratinės šaknies dėsnis 248
Kvartiliai 197 sek.; kvartilinis asimetrijos koeficientas 201 sek., dispersijos 203, 204

Lastelė 110
„ličba" 14
Lietuvos Metrika 14
Liniagrama 125
Lyginiai, ž. Statistiniai lyginiai

Masė 54, 55, 56, 59
Matematinė statistika 31
Mažiausiųjų kvadratų metodas 329 sek.
Medianas 107, 157, 169, 172 sek., 177, 189
Metodai, ž. Statistika
Mirtingumo lyginiai 105, 140, 142—145, grynieji 147—153
Mirtingumo lentelės 150, 154
missi 13
Modus 157

Natūrinio prieauglio lyginiai 143, 144; grynieji 147
Nedalyvaujančioji masė 146, 147
Netiesioginis stebėjimo metodas 67 sek.
Netto santykio lyginiai, ž. Statistiniai lyginiai
Normalinė dispersija, ž. Dispersija
Normalinė amžiaus grupių sudėtis 149—152
Nulis 112
Nuolatinė statistikos komisija 36, 37
Nusikaltimų lyginiai 144
Nusikaltimų biudžetas 26, 27, 29

Pagrindinė linkmė, esmė 328, 329, 358; pirmaeilė linkmė 331 sek., antraeilė, trečiaeilė etc. linkmė 337, 338; skaičiavimo metodai: mažiausių kvadratų metodas 339 sek., kintamųjų vidurkių metodas 338 sek., kintamųjų parametrų metodas 341 sek., sumavimų metodas 347 sek.
Paklaidos, statistinio stebėjimo 77, 88; galimoji 245, 248, 262, 269; normali arba vidutinė 245, 250, 262, 269; patikimoji 8, 245, 248, 261, 269; patikimosios paklaidos konstanta 260; Gauss'o paklaidų kreivė 259; vidurkio patikimoji 263 sek.; kvadratingojo nukrypimo patikimoji 263 sek.; reprezentacinėj statistikoj 270 sek., 273 sek.; reprezentacinės tikimybės 272, 273; reprezentacinių vidurkių 279 sek.
Pamiatnaja knižka 34
Pareto kreivė 195
Pažymys 56, 57

Pirminė statistika 67
Popis Zemski 13, 14 sek.
Politinė aritmetika 24, 29, 30
Prasikaltimų biudžetas 26, 27, 29
Prasikaltėlių lyginiai 144
Prieskiltis 110
Privatinė statistika 32

Registracijos kortelė 57
Regresijos formula 321
Reprezentacinė statistika, matematinis pagrindimas 83—85; reprezentantai 55, atsitiktiniai 79, 80, tipiškieji 80, pranašumo bei pirmenybės klausimas 81, 82; paklaidos (žiūr.); reprezentantų tikslumas bei griežtumas 276—278; reprezentacinės statistikos dalinių tikimybė ir paklaidos 273 sek.; orientaciniai dydžiai 282 sek.; grupių sudarymas 283 sek.
Reprodukcijos koeficientai 153, 154

Santykiniai skaičiai 134 sek.
Sąrašų sistema 58, 73, 74, 86
Simetrija 193, 194
Skaitomoji diena 57, 59, 60, 62
Skaitomasis periodas 58, 62, 63, 64
Skaitomasis vienetas 56
Skiltis 110
Socialinė fizika 25, 27
Specifiniai lyginiai 143, 148, 154, 155
Statistika, žodžio kilmė 1, 21; apibrėžimas 2, 8; uždaviniai 4; istorija 9 sek.; aprašomoji 31, 81; matematinė 31; universitetinė 17 sek., 25; tiesioginė 67 sek.; netiesioginė 67 sek.; pirminė 67; antrinė 67; oficialinė 32; privatinė 32; stovio 59 sek.; reiškinių arba judėjimo 59; 62; reprezentacinė (žiūr.); kryptys 30, 31; metodas 30, 31, 53 sek.; konvencijos 34; tarptautiniai kongresai 46; tarptautiniai institutai 46, 47; šaltiniai 48 sek.: pirminiai 48, antriniai 48, 49, kompiliatyviniai 48, 49, Lietuvos 49, 50, tarptautiniai 50
Statistikos darbo organizacija 32 sek.; organai 32, 33; centralizacija 33, 36, 43; decentralizacija 31; Lietuvoje 34 sek.; Klaipėdos krašte 39 sek.
Statistinio stebėjimo metodai ir sistemos 66 sek.; tiesioginis 67 sek., netiesioginis 67 sek., individualinis 57, 73 sek., 86, kolektyvinis 57, 74 sek., 86, kortelių sistema 58, 73, 74, 86, sąrašų sistema 58, 73, 74, 86, ištisinis 53, 54, 71, dalinis 71, 81, 82, reprezentacinis (žiūr.), anketų 78, visuotinis 53, 54, 78, 81, 82
Statistinė lentelė 98, 102, 109 sek., antgalvis 109, brūkšnelis 112, ląstelė 110, normaliniai formatai 110, numeracija 111, nulis 112, prieskiltis 110, skiltis 110, skaičių eilutė 110, skaičių šriftas 111, specialiniai formatai 110, taškas 112, sudarymo pavyzdžiai 115 sek.
Statistiniai daliniai 135 sek., 138, proporciniai 135, formulos 136, 137
Statistiniai lygumai 135, 137 sek., bendrieji (brutto) 142—145, grynieji (netto) 142, 143, 145 sek., specifiniai 143, 148, 154, 155
Statistiniai dėsningumai 6, 7 sek., 23, 26, 27, 29, 234 sek., gimimų 6, 238, 239, mirimų 6, 23, 234—237, vedybų 235—237, prasikaltimų 26, 27, 29, atsitiktinių įvykių 240 sek.
Statistinis paralelizmas 304, 320
Statistinis prasmingumas ir tikslingumas 98, 99, 100, 102
Statistinės grupės 97, 113
Statistiniai vienetai 103
Statistiniai koeficientai 143
Statistinio analizio uždaviniai 133, 134
Statistinis pažymys 56, 57
Statistinis pavaizdumas 100
Statistinė žymė 56, 57
Statistinis vienetas 100

Statistinė eilutė 57
Statistikos Taryba 43
Statistikos komitetai (gubernijų) 34
Statistikos tarptautiniai institutai 46, 47
Statistikos tarptautiniai biurai 47
Stebėjimo metodai, žiūr. Statistiniai stebėjimo metodai
Stebimoji diena 57, 59, 60
Stebimasis laikas 65
Stirling'o formula 248, 249, 250
Stovio statistika 59 sek.
Surašomasis vienetas 56, žymė 57
Surašymo metodas ir sistemos, žiūr. Statistiniai stebėjimo metodai
Sumavimų metodas aritmet. vidurkiui ieškoti 165 sek., pagrindinei linkmei — 347 sek., kvadratingajam nukrypimai 230 sek.
Surašomasis lakštas 57, 86, 87, individualinis 57, kolektyvinis 57
Surašomoji diena 57, 61, 62, 65
Surašymai (Lietuvoje), Popis Zemski 14 sek., gyventojų 16, žemės ūkio 16, žiūr. gyventojų surašymai
Susikaupimai, žiūr. Dispersija

Šaltiniai (statistikos), žiūr. Statistika

Tarptautinis Statistikos Biuras 47
Tarptautinis Statistikos Institutas 46, 47, 109, 114
Tarptautinis Darbo Biuras 47
Tarptautiniai statistikos kongresai 46
Tarpt. Prekybos Rūmų Institutas 47
Tarpt. Žemės Ūkio Institutas 47
Tarpt. Pašto Sąjunga 47
Tarpt. Telegrafo Sąjunga 47
Tautų Sąjunga 47
Taškas 112
Tiesioginė statistika 67 sek.
Tiesioginis stebėjimas 67 sek.
Tikimybės, matematinė 245 sek., maksimalinė 245, 248, 249, 250, 254, kvadratinės šaknies dėsnis 248, Stirling'o formula 248—250, vidurinis kvadratingasis nukrypimas 250—253, Gauss'o paklaidų kreivė 259
Tikslinė statistika 95
Trijų sigmų taisyklė 101, 212, 213, 220, 243, 261, 269

Universitetinė statistika 17 sek., 25
Užsienių prekybos statistika 39, 65, 72

Valakinė žemės reforma 14
Vaisingumo koeficientai 140, 145, 147
Vidurkiai 157 sek., 185 sek., aritmetinis 157, 159 sek., 185 sek., suprastinti skaičiavimo metodai 161 sek., skaičiavimas iš dydžio grupių 169 sek., dominantas (žiūr.), medianas (žiūr.), matematiniai 182 sek., geometrinis 182, 186, 192, 398, harmoninis 183, 186, 192, antiharmoninis 183, 186, 192, kvadratinis 183, 185, 186, 192, jų tikslingas parinkimas 185 sek., patikimoji paklaida 263
Vidutinis žmogus 28, 30
Vidurinis kvadratingasis nukrypimas, žiūr. Dispersija
Visuotinis stebėjimo metodas 53, 54, 78, 81, 82

„Žagrė" 14
Žemės ūkio surašymas (Lietuvoje) 69
Žemės ūkio korespondentai 80
Žymė 56, 57

Sommaire.

Dans le présent ouvrage on s'est proposé d'étudier les questions de théorie et de méthodologie statistique. L'ouvrage est issu des leçons professées à la faculté de droit pendant quelques ans, et des travaux pratiques. Ainsi la publication actuelle offre un certain développement et les résultats obtenus par les recherches et les travaux pratiques scientifiques poursuivies en théorie et méthodologie statistique.

L'ouvrage en question a pour but non seulement d'aider aux élèves qui suivent le cours et les travaux pratiques, mais aussi à tous ceux qui désirent utiliser les ressources et les méthodes employées dans la statistique. C'est pourquoi on a introduit quelques questions générales peu communes avec la théorie et méthodologie statistique.

L'ouvrage est divisé en quatre parties. Dans la première, on commence, de prime abord, par donner quelques notions introductives, notamment en ce qui concerne la comparaison de statistique et de comptabilité, la science statistique et la vie sociale et économique, les régularités statistiques etc. On trouve ici un aperçu historique de recensements dans l'antiquité, au moyen âge, dans le temps modernes. On rapelle de même toute l'évolution historique de la science statistique jusqu'à nos jours. Le dernier chapitre de cette partie est consacré à l'organisation de travail de la statistique.

D'abord, les principes de l'oganisation du travail statistique en général ensuite l'organistation en Lithuanie, l'organisation internationale de statistique, puis, la littérature et les périodiques traitant de la statistique.

Dans la deuxième partie on insiste sur les questions générales de théorie et de méthodologie statistique (chap. A.), sur les conditions essentielles des observations statistiques: observation statistique de faits quantitatifs, observations statistiques des mouvements ou des phénomènes (chap. B.).

Après un exposé articulé des conditions essentielles de bonnes observations statistiques, le dernier chapitre de la deuxième partie étudie les méthodes et les systèmes de l'observation. Ainsi, dans ce chapitre sont étudées d'une manière étendue les observations et notations directes et indirectes, individuelles et collectives, la méthode représentative, l'organisation et le contrôle des observations, ainsi, que les autres problèmes concernant les observations statistiques.

Dans la troisième partie on insiste sur la représentation et le classement. Cette partie étudie d'abord l'essence et le but de représentation statistique, ensuite, les divers problèmes de classements et leur fondements méthodologiques. Le dernier chapitre de cette partie, s'occupe des procédés techniques (préparations des tableaux, titres, carac-

tère du tableau etc.). Enfin, on insiste sur la représentation graphique des observations statistiques.

La dernière partie de l'ouvrage aborde les méthodes de l'nalyse statistique, les formes et les procédés, les comparaisons simples et complexes, les moyennes statistiques, l'écart moyen et l'écart quadratique, divers types de distributions statistiques etc. Un chapitre est specialement consacré aux probabilités et régularités statistiques.

On trouve ici les probabilités statistiques et matématiques, erreurs probables ou moyennes dans les groupes statistiques, les probabilités et les erreus accidentielles dans la statistique représentative; on insiste sur l'interpolation, la corrélation etc.

Les derniers chapitres de la quatrième partie étudient l'ajoustement de groupes statistiques et les indices. Dans le dernier consacré aux indices, on insiste d'abord sur les catégories d'indices, puis sur les indices à base fixe, à base mobile, enfin sur les indices colectifs.

Outre des notions générales, des explications et des définitions de théorie et de méthodologie statistique, l'ouvrage comprend un nombre considérable des tableaux représentant les divers domaines de la vie sociale et économique.

VYTAUTO DIDŽIOJO UNIVERSITETO

Teisių Fakulteto leidiniai.

I.

Teisės Mokslų Biblioteka.

1. *v. Gareis Karl,* **Teisės enciklopedija ir metodologija.** Su prof. *P. Leono* prakalba. 1924 m., 65 p.
2. *Leonas Petras,* prof., **Teisės enciklopedijos paskaitos.** 2 leid. 1931 m. 263 p.
3. *Leonas P.,* prof., **Teisės filosofijos istorija.** 1928. VIII. + 429 p.
4. *Mačys Vl.,* prof., **Civilinio proceso paskaitos.** 1924 m., 448 p.
5. *Stankevičius V.,* doc., **Baudžiamoji teisė.** 1925 m., 292 p.
6. *Gumplowicz L.,* **Sociologijos pagrindai.** Vertė *A. Janulaitis.* 1929 m., XIX + 343 + VIII.
7. *Jaščenka A.,* **Tarptautinės teisės kursas,** I t. Konstitucinė tarptautinė teisė. 1931 m. XVIII + 723 p.
8. *Olivier-Martin,* **Prancūzų revoliucija ir korporatinė organizacija.** 1931. 16 p.
9. *Girard,* **Romėnų teisė,** I t. Vertė *A. Tamošaitis.* 1931 m. IV + 492 p.
10. *Römer'is M.,* prof., **Dabartinės konstitucijos,** antrasis rinkinys. 1932 m. 233 + 229 p. (Teisės istorijos įvadas 1—233 p.; II. Konstitucijų tekstai 1—229 p.).
11. **Lietuvos Žydų seimai XVII—XVIII amž.**
12. *Lappo J.,* **Lietuva ir Lenkija po 1569 m. Liublino unijos.** 1932 m. IV + 172 p.
13. *Brunner K.,* **Trumpa Vokietijos teisės istorija.**
14. *Lappo J.,* **Iš Vyriausiojo Lietuvos Tribunolo istorijos XVI amž.** 1932 m. 27 p.
15. *Beccaria C.,* **Apie nusikaltimus ir bausmes.** Išvertė *R. Burokas,* pratartis *Stankevičiaus.* 1935 m. XXVIII + 80 p.
16. *Lappo J.,* **Iš vyriausiųjų Lietuvos suvažiavimų istorijos XVI amž.** 1932 m. 43 p.
17. *Girard,* **Romėnų teisė,** II t. Vertė *A. Tamošaitis.* 1932 m. IV + 514 p.
18. *Kipp Th.,* **Romėnų teisės šaltinių istorija.**
19. *Girard P. F.,* **Romėnų teisė,** III t. Vertė *A. Tamošaitis.* 1933 m. IV + 514 p.
20. *Römer'is M.,* **Valstybė ir jos konstitucinė teisė,** I dalis. **Valstybė.** I tomas. 1934 m. XII + 340 p.
21. *Römer'is M.,* **Valstybė ir jos konstitucinė teisė,** I dalis. **Valstybė.** II tomas. 1935 m. XII + 412 p.
22. *Römer'is M.,* Lietuvos konstitucinės teisės paskaitos. I dalis. 1937 m. VIII + 527 p.
23. *Mestre Achille,* Iš viešosios prancūzų teisės. 1938 m. 25 p.
24. *Leonas P.,* **Sociologija.**
25. *A. Tamošaitis* ir *J. Kairys.* **Hammurabio Įstatymas** 1938 m. VIII+88.
26. *Römeris M.,* **Valstybė ir jos konstitucinė teisė,** II dalis. **Konstitucinės institucijos,** t. I. Suverenitetas. 1939. XII+272 p.

II.

Ekonominių Mokslų Biblioteka.

1. *Moravskis A.,* doc., **Lietuvos finansai,** I dalis. Mokslas apie valstybės išlaidas ryšy su Lietuvos biudžeto formavimusi ir vystymusi. 1925 m., 136 p.

2. *Moravskis A.*, doc., **Lietuvos finansai**, II dalis. 1925 m., 158 p.
3. *Pokrovskis N.*, prof., **Finansų mokslo pagrindai.** 1926 m., 522 p.
4. *Tugan-Baranovskis M.*, prof., **Politiškosios ekonomijos pagrindai.** Vertė doc. *P. Šalčius.* 1923 m., 329 p.
5. *Weber M.*, **Ūkio istorija.** Vertė *J. Pajaujis.* 1929 m., 308 p.
6. *Conrad J.*, **Politinės ekonomijos vadovėlis.** 1929 m., 188 p.
7. *Šalčius Petras*, **Kooperacija.** 1931 m., XII + 358 p.
8. *Eulenburg Franz*, **Ekonominė politika ir užsienio prekyba.** Vertė *J. Pajaujis.* 1931 m. XII + 487 p.
9. *Cassel Gustav*, **Teoretiškoji socialekonomija.** Vertė *P. Šalčius.* 1931 m.
10. *Oncken August*, **Politinės ekonomijos istorija.** Laikai prieš Adomą Smithą. 1932 m., 520 p.
11. *Gide Charles ir Rist Charles*, **Ekonominių teorijų istorija.** Nuo fiziokratų ligi mūsų dienų. Vertė *J. Pajaujis.* 1932 m., XVI + 776 p.
12. *Rimka A.*, **Socialekonominė statistika.** Teorija ir metodai, 1933 m. VIII + 572 p.
13. *Robertson D. H.*, **Pinigai.** Vertė *K. Lukša.* 1938 m. IV + 144 p.
14. *Rimka A.*, **Statistika**, teorija ir metodai. 1939. XVI+420 p.

III.

Teisių Fakulteto Darbai.

I tomas. 1924 m. red.: doc. *A. Rimka.*

1. *Leonas P.*, prof., **Kolektivizmo pradas Lietuvos civilinės teisės novelėse.**
2. *Mačys Vl.* — **Socialinis teisių mokslo tikslas ir mūsų žemės reformos įstatymas.**
3. *Rimka A.* — **Lietuvos visuomenės ūkio bruožai ligi Liublino unijos metu.**
4. *Šalčius P.*, doc. — **Kooperatyvų įstatymai;** 5. Kritika ir bibliografija; 6. Teisių Fakulteto apyskaita 1922/24 mokslo metams. 232 p.

II tomas. 1925/28 m. red.: prof. *A. Janulaitis.*

1. *Römer'is M.*, prof., **Trečiasis tarptautinis sociologijos kongresas Romoje.** 1—14 p.
2. *Leonas P.*, prof., **Dėl vieno įstatymų aiškinimo.** 15—22 p.
3. *Rimka A.*, doc., **Lietuvos prekybos santykiai ligi unijos su lenkais.** 23—86 p.
4. *Janulaitis A.*, prof. **Vyriausias Lietuvos Tribunolas XVI—XVII amž.** 1—172 p.
5. *Balogh'as El.*, **Teisinė pralaikėjimo palūkanų prigimtis.** 1—27 p.
6. *Gummerus H.*, **Piniginio ūkio kitimas senovės Romoje.** 1—15 p.

III tomas. 1927 m. red.: prof. *A. Janulaitis.*

1. *Römer'is M.*, prof., **Reprezentacija ir mandatas.** 1—312 p.
2. *Janulaitis A.*, prof., **Imperijos Rūmų Teismas.** 1—83 p.

IV tomas. 1928 m. red.: prof., *A. Janulaitis.*

1. *Janulaitis A.*, **Užnemunė po Prūsais** (1795—1807 m.). XII+404 p.
2. *Römer'is M.*, **Administracinis teismas.** XII + 359 p.

V tomas. 1929 m. red.: prof. *A. Janulaitis.*

1. *Šalčius P.*, doc., **Iš vertės ir kainų teorijų.** 56 p.
2. *Tamošaitis A.*, **Istoriškoji teisės mokykla Vokietijoj.** 179 p.
3. *Šalčius P.*, doc., **Kooperatizmas.** 135 p.
4. *Janulaitis A.*, prof., **Baudžiavų panaikinimo sumanymai Lietuvoje** (1817—1819 m.). 112 p.
5. *Veriga-Darevskis*, **Lietuvos mokesčiai XV—XVI amž.** 52 p.
6. *Janulaitis A.*, pro[illegible] p.